

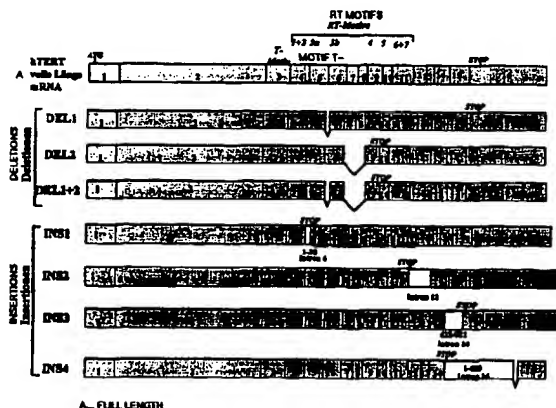
PCT
 WELTORGANISATION FÜR GEISTIGES EIGENTUM
 Internationales Büro
 INTERNATIONALE ANMELDUNG. VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
 INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)



<p>(51) Internationale Patentklassifikation ⁶ : C12N 15/54, 9/12, 15/11, 15/85, C12Q 1/68, A01K 67/027</p>	A2	<p>(11) Internationale Veröffentlichungsnummer: WO 99/33998</p> <p>(43) Internationales Veröffentlichungsdatum: 8. Juli 1999 (08.07.99)</p>
<p>(21) Internationales Aktenzeichen: PCT/EP98/08216</p> <p>(22) Internationales Anmeldedatum: 22. Dezember 1998 (22.12.98)</p> <p>(30) Prioritätsdaten: 197 57 984.1 24. Dezember 1997 (24.12.97) DE</p> <p>(71) Anmelder (für alle Bestimmungsstaaten ausser US): BAYER AKTIENGESELLSCHAFT [DE/DE]; D-51368 Leverkusen (DE).</p> <p>(72) Erfinder; und (75) Erfinder/Anmelder (nur für US): HAGEN, Gustav [DE/DE]; Bertha-von-Suttner-Strasse 31, D-51373 Leverkusen (DE). WICK, Maresa [DE/DE]; Andreas-Gryphius-Strasse 26, D-51065 Köln (DE). ZUBOV, Dmitry [RU/DE]; Roggen-dorfstrasse 59, D-51061 Köln (DE).</p> <p>(74) Gemeinsamer Vertreter: BAYER AKTIENGE-SELLSCHAFT; D-51368 Leverkusen (DE).</p>	<p>(81) Bestimmungsstaaten: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Veröffentlicht <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i></p>	

(54) Title: **REGULATORY DNA SEQUENCES OF THE HUMAN CATALYTIC TELOMERASE SUB-UNIT GENE, DIAGNOSTIC AND THERAPEUTIC USE THEREOF**

(54) Bezeichnung: **REGULATORISCHE DNA-SEQUENZEN DES GENS DER HUMANEN KATALYTISCHEN TELOMERASE-UNTEREINHEIT UND DEREN DIAGNOSTISCHE UND THERAPEUTISCHE VERWENDUNG**



(57) Abstract

The present invention relates to regulatory DNA sequences containing promotor sequences, in addition to intervening sequences, for the human catalytic telomerase sub-unit gene. The invention also relates to the use of said DNA sequences for pharmaceutical, diagnostic and therapeutic purposes, especially in the treatment of cancer and ageing.

(57) Zusammenfassung

Diese Erfindung betrifft regulatorische DNA-Sequenzen, beinhaltend Promotorsequenzen, sowie Intronsequenzen, für das Gen der humanen kalytischen Telomerase-Untereinheit. Darüber hinaus betrifft diese Erfindung die Verwendung dieser DNA-Sequenzen für pharmazeutische, diagnostische und therapeutische Zwecke, vor allem in der Behandlung von Krebs und Alterung.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

Regulatorische DNA-Sequenzen des Gens der humanen katalytischen
Telomerase-Untereinheit und deren diagnostische und therapeutische Verwen-
dung

5 Aufbau und Funktion der Chromosomenenden

- Das genetische Material eukaryontischer Zellen ist auf linearen Chromosomen verteilt. Die Enden der Erbanlagen werden, abgeleitet von den griechischen Wörtern *telos* (Ende) und *meros* (Teil, Segment), als Telomere bezeichnet. Die meisten
- 10 Telomere bestehen aus Wiederholungen von kurzen Sequenzen, die überwiegend aus Thymin und Guanin aufgebaut sind (Zakian, 1995). In allen bislang untersuchten Wirbeltieren werden die Telomere aus der Sequenz TTAGGG aufgebaut (Meyne *et al.*, 1989).
- 15 Die Telomere üben verschiedene wichtige Funktionen aus. Sie verhindern die Fusion von Chromosomen (McClintock, 1941) und damit die Entstehung von dizentrischen Erbanlagen. Solche Chromosomen mit zwei Centromeren können durch Verlust der Heterozygotie bzw. Verdopplung oder Verlust von Genen zur Entwicklung von Krebs führen.
- 20 Desweiteren dienen Telomere dazu, intakte Erbanlagen von beschädigten zu unterscheiden. So stellten Hefezellen ihre Zellteilung ein, wenn sie ein Chromosom ohne Telomer enthielten (Sandell und Zakian, 1993).
- 25 Eine weitere wichtige Aufgabe erfüllen Telomere bei der DNA-Replikation eukaryontischer Zellen. Im Gegensatz zu den zirkulären Genomen von Prokaryonten können die linearen Chromosomen der Eukaryonten von dem DNA Polymerase-Komplex nicht vollständig repliziert werden. Zur Initiation der DNA-Replikation sind RNA-Primer notwendig. Nach Abspaltung der RNA-Primer, Verlängerung der
- 30 Okazaki-Fragmente und anschließender Ligation fehlt dem neu-synthetisierten DNA-Strang das 5'-Ende, denn dort kann der RNA-Primer nicht durch DNA ersetzt

werden. Ohne besondere Schutzmechanismen würden daher die Chromosomen mit jeder Zellteilung schrumpfen ("end-replication problem"; Harley *et al.*, 1990). Die nicht-kodierenden Telomersequenzen stellen vermutlich eine Pufferzone dar, um dem Verlust von Genen vorzubeugen (Sandell und Zakian, 1993).

5

Darüberhinaus spielen Telomere auch eine wichtige Rolle bei der Regulation der zellulären Alterung (Olovnikov, 1973). Humane somatische Zellen zeigen in Kultur eine limitierte Replikationskapazität; sie werden nach einer gewissen Zeit seneszent. In diesem Zustand teilen sich die Zellen selbst nach Stimulierung mit Wachstumsfaktoren nicht mehr, sterben aber nicht, sondern bleiben metabolisch aktiv (Goldstein, 1990). Verschiedene Beobachtungen sprechen für die Hypothese, daß eine Zelle anhand der Länge ihrer Telomere bestimmt, wie oft sie sich noch teilen kann (Allsopp *et al.*, 1992).

10

15

Zusammenfassend besitzen die Telomere somit zentrale Funktionen bei der Alterung von Zellen sowie der Stabilisierung des genetischen Materials und Verhinderung von Krebs.

Das Enzym Telomerase synthetisiert die Telomere

20

Wie oben beschrieben können Organismen mit linearen Chromosomen ohne einen speziellen Schutzmechanismus ihr Genom nur unvollständig replizieren. Die meisten Eukaryonten verwenden zur Regeneration der Telomersequenzen ein spezielles Enzym, die Telomerase. In den bislang untersuchten Einzellern wird Telomerase konstitutiv exprimiert. Dagegen wurde in Menschen die Telomerase-Aktivität nur in Keimzellen und Tumorzellen gemessen, wogegen benachbartes somatisches Gewebe keine Telomerase enthielt (Kim *et al.*, 1994).

25

30

Funktionell kann die Telomerase auch als terminale Telomertransferase bezeichnet werden, die als Multiproteinkomplex im Zellkern lokalisiert ist. Während der RNA-Anteil der humanen Telomerase schon seit längerem bekannt ist (Feng *et al.*, 1995),

wurde kürzlich die katalytische Untereinheit dieser Enzymgruppe in verschiedenen Organismen identifiziert (Lingner *et al.*, 1997; vgl. unsere ebenfalls anhängige Anmeldung PCT EP/98/03468). Diese katalytischen Untereinheiten der Telomerase sind sowohl untereinander als auch zu bisher allen bekannten reversen Transkriptasen
5 auffällig homolog.

Auch in WO 98/14592 werden Nukleinsäure- und Aminosäuresequenzen der katalytischen Telomerase-Untereinheit beschrieben.

10 Aktivierung der Telomerase in menschlichen Tumoren

Eine Aktivität der Telomerase konnte in Menschen ursprünglich nur in Keimbahnzellen, nicht aber in normalen somatischen Zellen (Hastie *et al.*, 1990; Kim *et al.*, 1994) nachgewiesen werden. Nach der Entwicklung eines sensitiveren Nachweisverfahrens
15 (Kim *et al.*, 1994) wurde auch in hematopoietischen Zellen eine geringe Telomeraseaktivität detektiert (Broccoli *et al.*, 1995; Counter *et al.*, 1995; Hiyama *et al.*, 1995). Allerdings wiesen diese Zellen trotzdem eine Reduktion der Telomere auf (Vaziri *et al.*, 1994; Counter *et al.*, 1995). Noch ist nicht geklärt, ob die Menge an Enzym in diesen Zellen nicht ausreichend für eine Kompensation des Telomerverlustes ist, oder
20 ob die gemessene Telomerase-Aktivität von einer Subpopulation, z.B. unvollständig ausdifferenzierten CD34⁺38⁺-Vorläuferzellen, herrührt (Hiyama *et al.*, 1995). Zur Klärung wäre ein Nachweis der Telomerase-Aktivität in einer einzelnen Zelle nötig.

Interessanterweise wurde jedoch in einer großen Zahl der bislang getesteten Tumorgewebe eine signifikante Telomerase-Aktivität nachgewiesen (1734/2031, 85 %; Shay, 1997), während in normalem somatischem Gewebe keine Aktivität gefunden wurde (1/196, <1 %, Shay, 1997). Verschiedene Untersuchungen zeigten außerdem, daß in seneszenten Zellen, die mit viralen Oncoproteinen transformiert wurden, die Telomere weiterhin schrumpften und Telomerase nur in der Subpopulation entdeckt
25 werden konnte, die die Wachstumskrise überlebte (Counter *et al.*, 1992). In diesen immortalisierten Zellen waren auch die Telomere stabil (Counter *et al.*, 1992). Ähnli-

che Befunde aus Untersuchungen an Mäusen (Blasco *et al.*, 1996) stützen die Annahme, daß eine Reaktivierung der Telomerase ein spätes Ereignis in der Tumorgenese ist.

5 Basierend auf diesen Ergebnissen wurde eine "Telomerase-Hypothese" entwickelt, die den Verlust von Telomersequenzen und Zellalterung mit der Aktivität von Telomerase und der Entstehung von Krebs verbindet. In langlebigen Spezies wie dem Menschen kann das Schrumpfen der Telomere als ein Mechanismus zur Tumorsuppression angesehen werden. Ausdifferenzierte Zellen, die keine Telomerase
10 enthalten, stellen bei einer bestimmten Länge der Telomere ihre Zellteilung ein. Mutiert eine solche Zelle, so kann aus ihr nur dann ein Tumor entstehen, wenn die Zelle ihre Telomere verlängern kann. Ansonsten würde die Zelle weiterhin Telomersequenzen verlieren, bis ihre Chromosomen instabil werden und sie schließlich zugrunde geht. Die Reaktivierung der Telomerase ist vermutlich der Hauptmechanismus von Tumorzellen zur Stabilisation ihrer Telomere.
15

Aus diesen Beobachtungen und Überlegungen ergibt sich, daß eine Inhibition der Telomerase eine Therapie von Tumoren erlauben sollte. Konventionelle Krebstherapien mit Zytostatika oder kurzweiligen Strahlen schädigen nicht nur die Tumorzellen,
20 sondern alle sich teilenden Zellen des Körpers. Da aber außer Tumorzellen nur Keimbahnzellen eine signifikante Telomerase-Aktivität enthalten, würden Telomerase-Inhibitoren spezifischer die Tumorzellen angreifen und somit weniger unerwünschte Nebenwirkungen hervorrufen. In allen bislang getesteten Tumorgeweben wurde eine Telomerase-Aktivität nachgewiesen, so daß diese Therapeutika gegen alle
25 Krebsarten eingesetzt werden könnten. Die Wirkung von Telomerase-Inhibitoren würde dann eintreten, wenn die Telomere der Zellen sich soweit verkürzt haben, daß das Genom instabil wird. Da Tumorzellen meist kürzere Telomere aufweisen als normale somatische Zellen, würden zuerst Krebszellen durch Telomerase-Inhibitoren eliminiert werden. Zellen mit langen Telomeren, wie die Keimzellen, würden
30 dagegen erst viel später geschädigt werden. Telomerase-Inhibitoren stellen somit einen zukunftsweisenden Weg für die Therapie von Krebs dar.

Eindeutige Antworten auf die Frage nach der Art und den Angriffspunkten physiologischer Telomerase-Inhibitoren werden möglich sein, wenn auch die Regulation der Genexpression der Telomerase identifiziert ist.

5

Regulation der Genexpression in Eukaryonten

Die eukaryotische Genexpression, d.h. der zelluläre Informationsfluß von der DNA über die RNA zum Protein, weist vielfältige Ansatzpunkte für regulatorische Mechanismen auf. Einzelne Kontrollstufen sind z.B. die Gen-Amplifikation, Rekombination von Genloci, Chromatinstruktur, DNA-Methylierung, Transkription, posttranskriptionelle mRNA-Modifikationen, mRNA-Transport, Translation und post-translationale Proteinmodifikationen. Nach bisherigen Studien besitzt die Kontrolle auf der Ebene der Transkriptionsinitiation die größte Bedeutung (Latchman, 1991).

15

Unmittelbar stromaufwärts vom Transkriptionsstart eines von der RNA-Polymerase II transkribierten Gens liegt eine Region, die für die Steuerung der Transkription verantwortlich ist und als Promotorregion bezeichnet wird. Ein Vergleich der Nukleotidsequenzen von Promotorregionen vieler bekannter Gene zeigt, daß bestimmte Sequenzmotive in dieser Region häufig vorkommen. Zu diesen Elementen gehören unter anderem die TATA-Box, die CCAAT-Box und die GC-Box, die von spezifischen Proteinen erkannt werden. Die TATA-Box, die etwa 30 Nukleotide stromaufwärts vom Transkriptionsstart entfernt positioniert ist, wird z.B. von der TFIID-Untereinheit TBP („TATA-box binding protein“) erkannt, wogegen bestimmte GC-reiche Sequenzabschnitte vom Transkriptionsfaktor Sp1 („specificity protein1“) spezifisch gebunden werden.

25

Funktionell kann man den Promotor in einen regulativen und einen konstitutiven Abschnitt unterteilen (Latchman, 1991). Der konstitutive Kontrollbereich umfaßt den sogenannten Kernpromotor („corepromoter“), der die korrekte Initiation der Transkription ermöglicht. Er enthält die als UPE's (upstream promoter elements“) be-

30

schriebenen Sequenzelemente, die für eine effiziente Transkription notwendig sind. Die regulativen Kontrollabschnitte, die mit den UPE's verflochten sein können, weisen Sequenzelemente auf, die an der signalabhängigen Regulation der Transkription durch Hormone, Wachstumsfaktoren usw. beteiligt sein können. Sie vermitteln gewebs- oder zellspezifische Promotoreigenschaften.

Ein charakteristisches Merkmal eukaryotischer Gene sind DNA-Abschnitte, die über vergleichsweise große Distanzen hinweg Einfluß auf die Genexpression nehmen können. Diese Elemente können stromaufwärts, stromabwärts oder innerhalb einer Transkriptionseinheit lokalisiert sein und unabhängig von ihrer Orientierung ihre Funktion wahrnehmen. Diese Sequenzabschnitte können die Promotoraktivität verstärken (Enhancer) oder abschwächen (Silencer). Ähnlich wie die Promotorregionen beherbergen auch Enhancer und Silencer mehrere Bindungsstellen für Transkriptionsfaktoren.

Die Erfindung betrifft die DNA-Sequenzen aus der 5'-flankierenden Region des Gens der katalytisch aktiven humanen Telomerase-Untereinheit sowie Intron-Sequenzen für dieses Gen.

Die Erfindung betrifft insbesondere die 5'-flankierende regulatorische DNA-Sequenz, enthaltend die Promotor-DNA-Sequenz für das Gen der humanen katalytischen Telomerase Untereinheit gemäß Fig. 10 (SEQ ID NO 3).

Die Erfindung betrifft weiterhin regulatorisch wirksame Teilbereiche der 5'-flankierenden regulatorischen DNA-Sequenz gemäß Fig. 4 (SEQ ID NO 1).

Weiterhin sind Gegenstand der vorliegenden Erfindung Intron-Sequenzen für das Gen der humanen katalytischen Telomerase-Untereinheit, insbesondere solche, die regulatorische Wirkung haben. Die erfindungsgemäßen Intronsequenzen werden im Rahmen von Beispiel 5 detailliert beschrieben (vgl. SEQ ID NO 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 und 20).

Die Erfindung betrifft weiterhin ein rekombinantes Konstrukt, das die erfindungsgemäßen DNA-Sequenzen, insbesondere die 5'-flankierende DNA-Sequenz des Gens der humanen katalytischen Telomerase Untereinheit oder Teilbereiche davon
5 beinhaltet.

Bevorzugt sind rekombinante Konstrukte, die neben den erfindungsgemäßen DNA-Sequenzen, insbesondere der 5'-flankierenden DNA-Sequenz des Gens der humanen katalytischen Telomerase Untereinheit oder Teilbereichen davon, eine oder mehrere
10 weitere DNA-Sequenzen, die für Polypeptide oder Proteine kodieren, enthalten.

Gemäß einer besonders bevorzugten Ausführungsform kodieren diese weiteren DNA-Sequenzen für antitumorale Proteine.

15 Besonders bevorzugte antitumorale Proteine sind solche, die die Angiogenese direkt oder indirekt inhibieren. Zu diesen Proteinen zählen beispielsweise:

Plasminogenaktivatorinhibitor (PAI-1), PAI-2, PAI-3, Angiostatin, Endostatin, Platelet factor 4, TIMP-1, TIMP-2, TIMP-3, Leukemia Inhibitory Factor (LIF).
20

Ebenfalls besonders bevorzugt sind antitumorale Proteine, welche direkt oder indirekt eine zytostatische Wirkung auf Tumoren aufweisen. Hierzu zählen im besonderen:

25 Perforin, Granzym, IL-2, IL-4, IL-12, Interferone, wie beispielsweise IFN- α , IFN- β , IFN- γ , TNF, TNF- α , TNF- β , Oncostatin M; Tumorsuppressorgene, wie z.B. p53, Retinoblastoma.

Weiterhin besonders bevorzugt sind antitumorale Proteine, welche gegebenenfalls
30 zusätzlich zur antitumoralen Wirkung Entzündungen stimulieren und hierdurch zur Elimination von Tumorzellen beitragen. Hierzu zählen beispielsweise:

5 RANTES, Monocyte chemotactic and activating factor (MCAF), IL-8, Macrophage inflammatory protein (MIP-1 α , - β), Neutrophil activating protein-2 (NAP-2), IL-3, IL-5, human leukemia inhibitory factor (LIF), IL-7, IL-11, IL-13, GM-CSF, G-CSF, M-CSF.

10 Weiterhin besonders bevorzugt sind antitumorale Proteine, welche aufgrund ihrer Wirkung als Enzyme in der Lage sind, Vorstufen eines antitumoralen Wirkstoffes in einen antitumoralen Wirkstoff zu überführen. Zu diesen Enzymen zählen beispielsweise:

15 Herpes Simplex Virus Thymidinkinase, Varizella Zoster Virus Thymidinkinase, bakterielle Nitroreductase, bakterielle β -Glukuronidase, pflanzliche β -Glukuronidase aus *Secale cereale*, humane Glukuronidase, humane Carboxypeptidase, bakterielle Carboxypeptidase, bakterielle β -Lactamase, bakterielle Cytosindeaminidase, humane Katalase bzw. Phosphatase, humane alkalische Phosphatase, Typ 5 saure Phosphatase, humane Lysooxidase, humane saure D-Aminooxidase, humane Glutathion Peroxidase, humane Eosinophilen Peroxidase, humane Schilddrüsen Peroxidase.

20 Die obengenannten rekombinanten Konstrukte können auch DNA-Sequenzen enthalten, die für Faktor VIII, IX oder Teilfragmente davon kodieren. Zu diesen DNA-Sequenzen zählen auch andere Blutgerinnungsfaktoren

25 Die obengenannten rekombinanten Konstrukte können auch DNA-Sequenzen enthalten, die für ein Reporterprotein kodieren. Zu diesen Reporterproteinen zählen beispielsweise:

30 Chloramphenicolacetyltransferase (CAT), Glühwürmchen Luziferase (LUC), β -Galaktosidase (β -Gal), Sezernierte alkalische Phosphatase (SEAP), Humanes Wachstumshormon (hGH), β -Glukuronidase (GUS), Grün-fluoreszierendes Protein (GFP) und alle davon abgeleiteten Varianten, Aquarin, Obelin.

Erfindungsgemäße rekombinante Konstrukte können auch DNA kodierend für die humane katalytische Telomerase Untereinheit und deren Varianten und Fragmente in antisense Orientierung enthalten. Gegebenenfalls können diese Konstrukte auch
5 andere Protein-Untereinheiten der humanen Telomerase und die Telomerase-RNA-Komponente in antisense Orientierung enthalten.

Die rekombinanten Konstrukte können neben der DNA, kodierend für die humane katalytische Telomerase Untereinheit, sowie deren Varianten und Fragmente auch
10 andere Protein-Untereinheiten der humanen Telomerase und die Telomerase-RNA-Komponente enthalten.

Die Erfindung betrifft weiterhin einen Vektor, enthaltend die oben genannten erfindungsgemäßen DNA-Sequenzen, insbesondere die 5'-flankierenden DNA-Sequenzen,
15 zen, sowie eine oder mehrere der oben genannten anderen DNA-Sequenzen.

Bevorzugter Vektor für solche Konstrukte ist ein Virus, beispielsweise ein Retrovirus, Adenovirus, adeno-assoziiertes Virus, Herpes Simplex Virus, Vaccina Virus, lentivirales Virus, Sindbis Virus und ein Semliki Forest Virus.
20

Ebenfalls bevorzugt sind Plasmide als Vektoren.

Die Erfindung betrifft weiterhin pharmazeutische Präparate, enthaltend erfindungsgemäße rekombinante Konstrukte bzw. Vektoren; beispielsweise eine Zubereitung in einem kolloidalen Dispersionssystem.
25

Geeignete kolloidale Dispersionssysteme sind beispielsweise Liposome oder Polylysin-Liganden.

30 Die Zubereitungen der erfindungsgemäßen Konstrukte bzw. Vektoren in kolloidalen Dispersionssystemen können um einen Liganden ergänzt sein, der an Membranstruk-

turen von Tumorzellen bindet. Ein solcher Ligand kann z.B. an das Konstrukt bzw. den Vektor angeknüpft sein oder auch Bestandteil der Liposomenstruktur sein.

5 Geeignete Liganden sind insbesondere polyklonale oder monoklonale Antikörper oder Antikörperfragmente hiervon, die mit ihren variablen Domänen an Membranstrukturen von Tumorzellen binden, oder endständige Mannose-tragende Substanzen, Zytokine, Wachstumsfaktoren oder Fragmente bzw. Teilsequenzen hiervon, die an Rezeptoren auf Tumorzellen binden.

10 Entsprechende Membranstrukturen sind beispielsweise Rezeptoren für ein Zytokin oder einen Wachstumsfaktor, wie z.B. IL-1, EGF, PDGF, VEGF, TGF β , Insulin oder Insulin-like Growth Factor (IGF), oder Adhäsionsmoleküle, wie z. B. SLeX, LFA-1, MAC-1, LECAM-1 oder VLA-4, oder der Mannose-6-Phosphat-Rezeptor.

15 Zur vorliegenden Erfindung gehören pharmazeutische Zubereitungen, die neben den erfindungsgemäßen Vektorkonstrukten auch nichttoxische, inerte, pharmazeutisch geeignete Trägerstoffe enthalten können. Vorstellbar sind die Applikation (z.B. intravenös, intraarteriell, intramuskulär, subkutan, intradermal, anal, vaginal, nasal, transdermal, intraperitoneal, als Aerosol oder oral) am Ort eines Tumors oder die systemische Applikation dieser Zubereitungen.

20

Die erfindungsgemäßen Vektorkonstrukte können in der Gentherapie eingesetzt werden.

25 Die Erfindung betrifft weiterhin eine rekombinante Wirtszelle, insbesondere eine rekombinante eukaryotische Wirtszelle, enthaltend die vorstehend beschriebenen Konstrukte bzw. Vektoren.

Die Erfindung betrifft weiterhin ein Verfahren zur Identifizierung von Substanzen, die die Promotor-, Silencer- oder Enhanceraktivität der katalytischen Telomerase Untereinheit beeinflussen, wobei dieses Verfahren folgende Schritte umfaßt:

30

A. Zugabe einer Kandidatensubstanz zu einer Wirtszelle, enthaltend die erfindungsgemäße regulatorische DNA-Sequenz, insbesondere die 5'-flankierende regulatorische DNA-Sequenz für das Gen der humanen katalytischen
5 Telomerase-Untereinheit oder einen regulatorisch wirksamen Teilbereich davon, funktionell verknüpft mit einem Reportergen,

B. Messung des Substanzeffektes auf die Reportergenexpression.

10 Das Verfahren kann eingesetzt werden zur Identifizierung von Substanzen, die die Promotor-, Silencer- oder Enhanceraktivität der katalytischen Telomerase Untereinheit verstärken.

Das Verfahren kann weiterhin eingesetzt werden zur Identifizierung von Substanzen,
15 die die Promotor-, Silencer- oder Enhanceraktivität der katalytischen Telomerase Untereinheit inhibieren.

Die Erfindung betrifft weiterhin ein Verfahren zur Identifizierung von Faktoren, die spezifisch an Fragmente der erfindungsgemäßen DNA-Fragmente, insbesondere der
20 5'-flankierenden regulatorischen DNA-Sequenz der katalytischen Telomerase Untereinheit, binden. Diese Methode beinhaltet ein Screening einer Expressions-cDNA-Bibliothek mit der vorstehend beschriebenen DNA-Sequenz oder Teilfragmenten unterschiedlichster Länge als Sonde.

25 Die vorstehend beschriebenen Konstrukte bzw. Vektoren können auch zur Herstellung transgener Tiere verwendet werden.

Die Erfindung betrifft weiterhin ein Verfahren zur Detektion Telomerase-assoziiierter Zustände bei einem Patienten, das folgende Schritte umfaßt:

30

- A. Inkubation eines Konstruktes bzw. Vektors, enthaltend die erfindungsgemäße DNA-Sequenz, insbesondere die 5'-flankierende regulatorische DNA-Sequenz für das Gen der humanen katalytischen Telomerase-Untereinheit oder einen regulatorisch wirksamen Teilbereich davon sowie ein Reportergen mit Körperflüssigkeiten oder zellulären Proben,
- B. Detektion der Reportergenaktivität, um einen diagnostischen Wert zu erhalten;
- C. Vergleich des diagnostischen Werts mit Standardwerten für das Reportergenkonstrukt in standardisierten normalen Zellen oder Körperflüssigkeiten des gleichen Typs wie die Testprobe;

Detektion diagnostischer Werte, die höher oder niedriger als Standardvergleichswerte liegen, indiziert einen Telomerase-assoziierten Zustand, der wiederum einen pathogenen Zustand indiziert.

Erläuterung der Abbildungen:

Fig. 1: Southern Blot-Analyse mit genomischer DNA verschiedener Spezies

A: Foto eines Ethidiumbromid gefärbten 0,7 %igen Agarosegels mit etwa 4 µg Eco RI geschnittener genomischer DNA. Die Spur 1 enthält Hind III geschnittene λ-DNA als Größenmarker (23,5, 9,4, 6,7, 4,4, 2,3, 2,0, und 0,6 kb). Die Spuren 2 bis 10 enthalten genomische DNA von Mensch, Rhesusaffe, Sprague Dawley Ratte, BALB/c Maus, Hund, Rind, Kaninchen, Huhn und Hefe (*Saccharomyces cerevisiae*).

B: Zu Fig.1 A korrespondierendes Autoradiogramm einer Southern Blot-Analyse, hybridisiert mit einer radioaktiv-markierten etwa 720 bp langen hTC-cDNA Sonde.

Fig. 2: Restriktionsanalyse der rekombinanten λ -DNA des Phagenklons P12, der mit einer Sonde aus dem 5'-Bereich der hTC-cDNA hybridisiert.

5 Die Abbildung zeigt ein Foto eines Ethidiumbromid gefärbten 0,4 %igen Agarosegels. Die Spuren 1 und 2 enthalten Eco RI/Hind III geschnittene λ -DNA bzw eine 1 kb Leiter der Firma Gibco als Größenmarker. Die Spuren 3 - 7 enthalten 250 ng mit Bam HI (Spur 3), Eco RI (Spur 4), Sal I (Spur 5), Xho I (Spur 6) und Sac I (Spur 7) geschnittene DNA des
10 rekombinanten Phagens. Die Pfeile kennzeichnen die zwei λ -Arme des Vektors EMBL3 Sp6/T7.

Fig. 3: Restriktionsanalyse und Southern Blot-Analyse der rekombinanten λ -DNA des Phagenklons, der mit einer Sonde aus dem 5'-Bereich der hTC-cDNA hybridisiert.

15 A: Die Abbildung zeigt ein Foto eines Ethidiumbromid gefärbten 0,8%igen Agarosegels. Die Spuren 1 und 15 enthalten eine 1 kb Leiter der Firma Gibco als Größenmarker. Die Spuren 2 bis 14 enthalten 250 ng geschnittene λ -DNA vom rekombinanten Phagenklon. Als Enzyme wurden eingesetzt: Spur 2: Sac I, Spur 3: Xho I, Spur 4: Xho I, Xba I, Spur 5: Sac I, Xho I, Spur 6: Sal I, Xho I, Xba I, Spur 7: Sac I, Xho I, Xba I, Spur 8: Sac I, Sal I, Xba I, Spur 9: Sac I, Sal I, BamH I, Spur 10: Sac I, Sal I, Xho I, Spur 11: Not I, Spur 12: Sma I, Spur 13: leer, Spur
20 14: nicht verdaut.

B: Zu Fig.3 A korrespondierendes Autoradiogramm einer Southern Blot-Analyse. Als Sonde für die Hybridisierung wurde ein etwa 420 bp langes 5'-hTC-cDNA Fragment eingesetzt.

30

Fig. 4: Partielle DNA-Sequenz der 5'-flankierenden Region und des Promotors vom Gen der humanen katalytischen Telomerase-Untereinheit. Das ATG-Startcodon ist in der Sequenz fett hervorgehoben. Die dargestellte Sequenz entspricht SEQ ID NO 1.

5

Fig. 5: Identifizierung des Transkriptionsstarts durch Primer Extension-Analyse.

10

Die Abbildung zeigt ein Autoradiogramm eines denaturierenden Polyacrylamidgels, welches zur Darstellung einer Primer Extension-Analyse gewählt wurde. Als Primer wurde ein Oligonukleotid mit der Sequenz 5'GTTAAGTTGTAGCTTACACTGGTTCTC 3' benutzt. In der Spur 1 wurde die Primer Extension Reaktion aufgetragen. Die Spuren G, A, T, C, stellen die Sequenzreaktionen mit dem gleichen Primer und den entsprechenden Dideoxynukleotiden dar. Der fette Pfeil kennzeichnet den Haupt-Transkriptionsstart, die dünnen Pfeile weisen auf drei Neben-Transkriptionsstartpunkte hin.

15

20

Fig. 6: cDNA Sequenz der humanen katalytischen Telomerase-Untereinheit (hTC; vgl. unsere anhängige Anmeldung PCT/EP/98/03468). Die dargestellte Sequenz entspricht SEQ ID NO 2.

Fig. 7: Strukturelle Organisation und Restriktionsmappe des humanen hTC-Gens und dessen 5'- und 3'-flankierende Region.

25

30

Exons sind als durchnummerierte schwarz ausgefüllte Rechtecke und Introns als nicht ausgefüllte Bereiche hervorgehoben. Nichttranslatierte Sequenzabschnitte in den Exons sind schraffiert. Die Translation startet in Exon 1 und endet in Exon 16. Restriktionsenzymschnittstellen sind wie folgt gekennzeichnet: S, SacI; X, XhoI. Die relative Anordnung der fünf Phagenklone (P2, P3, P5, P12, P17) und des Produktes aus dem „Genomic walking“ sind durch dünne Linien hervorgehoben. Wie durch

die Punkte gekennzeichnet, ist die Sequenz von Intron 16 nur teilweise entschlüsselt.

Fig. 8: HTL Splicevarianten.

5

10

15

A: Schematische Struktur der hTC mRNA Splicevarianten. Die vollständige hTC mRNA ist als grau unterlegtes Rechteck im oberen Bereich der Abb. dargestellt. Die 16 Exons sind entsprechend ihrer Größe dargestellt. Der Translationsstart (ATG) und das Stop-Codon, sowie das Telomerase-spezifische T-Motiv und die sieben RT-Motive sind hervorgehoben. Die hTC-Varianten sind in Deletions- und Insertionsvarianten unterteilt. In den Deletionen sind die fehlenden Exonsequenzen markiert. Die Insertionen sind durch zusätzliche weiße Rechtecke hervorgehoben. Größe und Herkunft der insertierten Sequenzen sind angegeben. Neu entstandene Stop-Codons sind markiert. Die Größe der Insertion von Variante INS2 ist unbekannt.

20

25

B: Exon Intron Übergänge der hTC-Splice-Varianten. Nichtgespligte 5'- und 3'-flankierte Sequenzen sind als weiße Rechtecke hervorgehoben. Die Herkunft der Exon und Intron Sequenzen ist angegeben. Intron und Exon Sequenzen sind in Kleinbuchstaben, bzw. in Großbuchstaben dargestellt. Die Donor und Akzeptor Sequenzen der Splicestellen sind als graue Rechtecke unterlegt und deren Exon Intron Herkunft ist ebenfalls angegeben.

Fig. 9: Identifizierung des Transkriptionsstarts durch RT-PCR Analyse.

30

Die RT-PCR wurde mit cDNA-Bibliothek aus HL 60 Zellen und genomischer DNA als Positivkontrolle durchgeführt. Ein gemeinsamer 3'-Primer hybridisiert an eine Sequenzregion aus Exon 1. Die Position der verschiedenen 5' Primer in der kodierenden Region oder der 5'-flankierenden Region ist angegeben. In der Negativkontrolle wurde keine

Template-DNA in der PCR-Reaktion zugegeben. M: DNA-Größenmarker.

Fig. 10: Nukleotidsequenz und Strukturmerkmale des hTC-Promotors.

5 11273 bp der 5'-flankierenden hTC Gensequenz, beginnend mit dem Translationsstartcodon ATG (+1) sind dargestellt. Die putative Region des Translationsstarts ist unterstrichen. Mögliche regulatorische Sequenzabschnitte innerhalb der 4000 bp stromaufwärts des Translationsstarts sind umrandet. Die dargestellte Sequenz entspricht SEQ ID NO 3.

10

Fig. 11: Aktivität des hTC-Promotors in HEK-293 Zellen.

Im oberen Bereich der Abbildung sind die ersten 5000 bp der 5'-flankierenden hTC Genregion schematisch dargestellt. Das ATG-Startcodon ist hervorgehoben. CpG reiche Inseln sind durch graue Rechtecke markiert. Auf der linken Abbildungsseite sind die Größen der hTC Promotor-Luziferase Konstrukte dargestellt. Das promotorlose pGL2-Basic Konstrukt und das SV40 Promotorkonstrukt pGL2-Pro wurden in jeder Transfektion als Kontrollen eingesetzt. Auf der rechten Abbildungsseite ist die relative Luziferaseaktivität der verschiedenen Promotorkonstrukte in HEK-Zellen als durchgehende Balken gezeigt. Die Standardabweichung ist angegeben. Die Zahlenwerte repräsentieren den Durchschnitt von zwei unabhängigen Experimenten, die in Duplikaten durchgeführt wurden.

15

20

25 Tab. 1: Exon Intron Übergänge des hTC-Gens

Aufgelistet sind die Nukleotidsequenzen an den 3'- und 5' Spliceübergängen des hTC-Gens. Die Konsensussequenzen für Donor und Akzeptorsequenzen (AG und GT) sind durch graue Rechtecke unterlegt. Intronsequenzen (Kleinbuchstaben) und Exonsequenzen (Großbuchstaben), die die Spliceakzeptor- und Donorstellen flankieren sind gezeigt. Die Größe der Exons und Introns ist in bp angegeben.

30

Tab. 2: Potentielle Bindungsstellen für DNA-bindende Faktoren in der Nukleotidsequenz von Intron 2

5 Die Suche nach möglichen DNA-bindenden Faktoren (z.B. Transkriptionsfaktoren) wurde mit dem „Find Pattern“-Algorithmus aus dem „GCG Sequenz Analysis“ Programmpaket der „Genetics Computer Group“ (Madison, USA) durchgeführt. Aufgelistet sind die Abkürzungen der identifizierten DNA-bindenden Faktoren und deren Lokalisation in
10 Intron 2.

Tab. 1

3' Acceptor Sequence			5' Donor Sequence					
Intron	Exon	Exon No.	bp	Exon	Intron	Intron on	bp	
						No.		No.
5' flanking Region								
caggcgcttcccccgag	GTTCAGGCAGCGCTGCGT	1	281	CGCCCCCTCCTTCGCCAG	gtggcctccccgggtcg	1	104	
catgtccttctegttttag	GTGTCCTGCTGAAGGAGC	2	1354	TGGCTGCGCAGGAGCCAG	gtgaggagggtgtggcgt	2	8616	
gagggtctctctattgag	GGGTTGGCTGTGTTCCGGC	3	196	TGCAAAAGCATTGGAAATCAG	gtactgtatccccacgcca	3	2089	
cccatgtgtcccccgag	ACAGCACTTGAAGAGGGTG	4	181	GTTCGCAGAGAAAAAGAGG	gtggtgtgtctttgttta	4	687	
ctcgctccactcacadag	GCCGAGGCTCTCACCCTGA	5	180	TGAGCTGTACTTTGTCAAG	gtgggtgccggggaacccc	5	494	
ccctctcctctgcgggtag	GTGGATGTGACGGGCGCGT	6	156	CAAGGCCTTCAAGAGCCAC	gtgaggttcaagtgtgata	6	>4660	18
ctcccgctgtgttttcgtag	GTCTCTACCTTGACAGACC	7	96	TGCCGTGCTCATCGAGCAG	gtctgggcactgcccctgca	7	980	
ctgtgtcttccccgccdag	AGTCCTCCTGCTGAATGAGG	8	86	CCGTGCGCATCAGGGGCAA	gtgagtcagggtggccaggt	8	2484	
gtattttcccttatttttag	GTCTACGTCCAGTGCCAG	9	114	CGGGGATTCGGGGGACGG	gtgagcctctcttcccc	9	1984	
cattgccccctctgccttag	GCTGCTCCTGCGTTTGGTG	10	72	ACGCGAAAACTTCTCTCAG	gtgagccccgtgccgtgtg	10	1871	
attccccctgtgtcttag	GACCTGCTCCGAGGTGTC	11	189	TGCAGAGCGACTACTCCAG	gtgagcgcaactggccgga	11	3804	
tctttattggcgactcttag	CTATGCCCGGACCTCCATC	12	127	CCTGTTTCTGATTTGCAG	gtgagcaggctgatgtca	12	880	
ctgtcgcgcactctcttag	GTGAACAGCCTCCAGACGG	13	62	TCCTGCTGCAGCGGTACAG	gtgagcgcaccacaagggg	13	3187	
agcctctgttttcccccdag	GTTCACGCATGTGTGCTG	14	125	CTGAAAGCCAAAGAACGAG	gtatgtcagggtgcttgcc	14	781	
tctgattttggcccccgtag	GGATGCTGCTGGGGGCCAA	15	138	CTGGGCTCACTCAGGACAG	gcaagtgtgggtggaggcc	15	536	
	CCCAGACGACGCTGAGTCG	16	664	TTTTTCAGTTTGTGAAAAA	3' flanking Region			

Tab. 2

Faktoren	Lokalisation in Intron 2
C/EBP	2925
CRE.2	2749
Sp1	2378, 4094, 4526, 4787, 4835, 4995
AP-2 CS3	5099
AP-2 CS4	2213, 3699, 4667, 5878, 5938, 6059, 6180, 6496
AP-2 CS5	5350, 5798, 5880, 5940, 6061, 6182, 6375, 6498
PEA3	934, 2505
P53	2125
GR uteroglobin	848, 1487, 2956
PR uteroglobin	3331
Zeste-white	1577, 1619, 1703, 1745, 1787, 1829, 1871, 1913, 1955, 1997, 2039, 2081, 3518, 3709, 4765, 5014, 5055
GRE	846
MyoD-MCK right site/rev	447, 509, 558, 1370, 1595, 1900, 2028, 2099, 4557
MyoD-MCK left site	108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030
Ets-1 CS	6408
API	3784, 4406
CREB	2801
GATA-1	839, 1390, 3154
c-Myc	108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030
CACCC site	991
CCAAT site	1224
CCAC box	992
CAAT site	463, 2395
Rb site	992, 4663
TATA	3650
CDEI	106, 1564, 1606, 1690, 1732, 1816, 1900, 1984

Beispiele

Das menschliche Gen für die katalytische Telomerase Untereinheit (ghTC), sowie die 5' und 3' liegenden Bereiche dieses Gens wurden kloniert, der Startpunkt der Transkription bestimmt, potentielle Bindungsstellen für DNA-bindende Proteine identifiziert, sowie aktive Promotorfragmente aufgezeigt. Die Sequenz der hTC-cDNA (Fig. 6) ist bereits in unserer ebenfalls anhängigen Anmeldung PCT/EP/98/03468 beschrieben. Wenn nicht gesondert erwähnt, beziehen sich sämtliche Angaben zur cDNA-Position auf diese Sequenz.

Beispiel 1

Durch eine genomische Southern Blot-Analyse wurde bestimmt, ob ghTC im menschlichen Genom ein Einzelgen darstellt oder mehrere Loci für das hTC-Gen bzw. eventuell auch ghTC-Pseudogene existieren.

Hierzu wurde ein kommerziell erhältlicher Zoo-Blot der Firma Clontech einer Southern Blot-Analyse unterzogen. Dieser Blot enthält 4 µg Eco RI geschnittene genomische DNA von neun verschiedenen Spezies (Mensch, Affe, Ratte, Maus, Hund, Rind, Kaninchen, Huhn und Hefe). Mit Ausnahme von Hefe, Huhn und Mensch wurde die DNA aus Nierengewebe isoliert. Die humane genomische DNA wurde aus Plazenta isoliert und die genomische DNA aus Huhn wurde aus Lebergewebe aufgereinigt. Im Autoradiogramm in Fig. 1 wurde als radioaktiv-markierte Sonde ein etwa 720 bp langes hTC-cDNA Fragment, isoliert aus der hTC cDNA, Variante Del2 (Position 1685 bis 2349 plus 2531 bis 2590 der Fig. 6 [Deletion 2; vergl. Beispiel 5 der Fig. 8]), eingesetzt. Die experimentellen Bedingungen für die Hybridisierung und die Waschschrte des Blots erfolgten in Anlehnung an Ausubel *et al.* (1987).

Im Fall der humanen DNA erkennt die Sonde zwei spezifische DNA-Fragmente. Das kleinere, etwa 1,5 bis 1,8 kb lange Eco RI-Fragment geht wahrscheinlich auf zwei

Eco RI-Schnittstellen in einem Intron der ghTC-DNA zurück. Aufgrund dieses Ergebnisses ist davon auszugehen, daß nur ein singuläres ghTC-Gen im menschlichen Genom vorliegt.

5 Beispiel 2

Zur Isolierung der 5' flankierenden hTC-Gensequenz wurden ca $1,5 \times 10^6$ Phagen einer humanen genomischen Plazenta-Genbibliothek (EMBL 3 SP6/T7 der Firma Clontech, Bestellnummer HL1067j) auf Nitrozellulosefilter (0,45 µm; Fa. Schleicher und Schuell) nach Angaben des Herstellers mit einem radioaktiv markierten, etwa
10 500 bp langen 5'-hTC-cDNA Fragment (Position 839 bis 1345 der Fig. 6) hybridisiert. Die Nitrozellulosefilter wurden zunächst in 2 x SSC (0,3 M NaCl; 0,5 M Tris-HCl, pH 8,0) und anschließend in einer Prähybridisierungslösung (50 % Formamid; 5 x SSPE, pH 7,4; 5 x Denhards-Lösung; 0,25 % SDS; 100 µg/ml
15 Heringsperma-DNA) zwei Stunden bei 42°C inkubiert. Für die Hybridisierung über Nacht wurde die Prähybridisierungslösung mit $1,5 \times 10^6$ cpm/ml Lösung denaturierter, radioaktiv markierter Probe ergänzt. Unspezifisch gebundene, radioaktive DNA wurde unter stringenten Bedingungen, d.h. durch drei fünfminütige Waschschr
20 itte mit 2 x SSC; 0,1 % SDS bei 55 bis 65 °C entfernt. Die Auswertung erfolgte durch Autoradiographie der Filter.

Die in dieser Primäruntersuchung identifizierten Phagenklone wurden aufgereinigt Ausubel *et al.* (1987). In weitergehenden Analysen stellte sich ein Phagenklon P12 als potentiell positiv heraus. Eine λ-DNA Präparation dieses Phagens Ausubel *et al.*
25 (1987) und der nachfolgende Restriktionsverdau mit Enzymen, die das genomische Insert in Fragmenten freisetzen, zeigte, daß dieser Phagenklon ein ca. 15 kb Insert im Vektor enthält (Fig. 2).

Zur Isolierung der vollständigen hTC-Gensequenz wurden in unabhängigen
30 Experimenten jeweils 1 bis $1,5 \times 10^6$ Phagen mit jeweils verschiedenen radioaktiv markierten Sonden wie oben beschrieben durchmustert.

Die in diesen Primäruntersuchungen identifizierten, für die entsprechenden Sonden positiven Phagenklone wurden aufgereinigt. Der Phagenklon P17 wurde mit einem etwa 250 bp langen hTC-cDNA Fragment (Position 1787 bis 2040 der Fig. 6) gefunden. Der Phagenklon P2 wurde mit einem etwa 740 bp langen hTC-cDNA Fragment (Position 1685 bis 2349 plus 2531 bis 2607 der Fig. 6 [Deletion 2; vergl. Beispiel 5]) identifiziert. Die Phagenklone P3 und P5 wurden mit einem 420 bp langen 3' hTC-cDNA Fragment (Position 3047 bis 3470 der Fig. 6) gefunden. Nach λ -DNA Präparation dieser Phagen und nachfolgendem Restriktionsverdau mit Enzymen, die das genomische Insert in Fragmenten freisetzen, wurden die Inserts in Plasmide umklont (Beispiel 4).

Beispiel 3

Um zu untersuchen, ob auch das 5'-Ende der hTC-cDNA im Insert des rekombinanten Phagenklons P12 vorliegt, wurde λ -DNA dieses Klons in einer Southern Blot Analyse mit einem radioaktiv markierten etwa 440 bp langen hTC-cDNA Fragment (Position 1 bis 440 der Fig. 6) aus dem extremen 5'-Bereich hybridisiert (Fig. 3).

Da die isolierte λ -DNA des positiven Klons auch mit dem extremen 5'-Ende der hTC-cDNA hybridisiert, enthält dieser Phage wahrscheinlich auch den das ATG-Startcodon flankierenden 5'-Sequenzbereich.

Beispiel 4

Um das gesamte 15 kb lange Insert des positiven Phagenklons P12 in Teilfragmenten umzuklonieren und anschließend zu sequenzieren, wurden zum DNA-Verdau Restriktionsendonukleasen ausgewählt, die zum einen das gesamte Insert aus EMBL3 Sp6/T7 freisetzen (vgl. Beispiel 2) und zusätzlich im Insert schneiden.

Insgesamt wurden ein etwa 8,3 und ein etwa 6,5 kb langes Xho I-Subfragment sowie ein etwa 8,5, ein etwa 3,5 und ein etwa 3 kb langes Sac I-Teilfragment in den Vektor pBluescript KS(+) (Fa. Stratagene) umkloniert. Durch Sequenzanalyse dieser Fragmente wurde die Nukleotidsequenz von 5123 bp 5'-flankierenden des ghTC-Genbereichs, ausgehend vom ATG-Startcodon bestimmt (Fig. 4; entsprechend SEQ ID NO 1). In der Fig. 4 sind die ersten (ausgehend vom ATG-Startcodon) 5123 bp dargestellt. In der Fig. 10 (entsprechend SEQ ID NO 3) die gesamte klonierte 5' Sequenz.

Um das gesamte ca. 14,6 kb große Insert des Phagenklons P17 in Teilfragmenten umzuklonieren, wurden zum DNA-Verdau Restriktionsendonukleasen ausgewählt, die zum einen das gesamte Insert aus EMLB3 Sp6/T7 freisetzen und zusätzlich einige Male im Insert schneiden. Durch Kombinationsverdau mit den Enzymen XhoI und BamHI wurden ein 7,1 kb, ein 4,2 kb und ein 1,5 kb großes XhoI-BamHI-Fragment sowie ein 1,8 kb großes BamHI-Fragment subkloniert. Der Kombinations-Restriktionsverdau mit den Enzymen XhoI und XbaI führte zur Klonierung von einem 6,5 kb großen XhoI-XbaI-Fragment, einem 6,5 kb und einem 1,5 kb großem XhoI-Fragment.

Die Umklonierung des ca. 17,9 kb großem Inserts des Phagenklons P2 in Subfragmente erfolgte durch Verdau mit dem Restriktionsenzym XhoI. Insgesamt wurde ein 7,5 kb, ein 6,4 kb sowie ein 1,6 kb langes XhoI-Subfragment kloniert. Durch Verdau mit dem Restriktionsenzym SacI wurde zusätzlich ein 4,8 kb, ein 3 kb, ein 2 kb sowie ein 1,8 kb großes SacI-Fragment subkloniert.

Das ca. 13,5 kb große Insert des Phagenklons P3 wurde durch Verdau mit den Restriktionsenzymen SacI bzw. XhoI subkloniert. Dabei wurden ein 3,2 kb, ein 2 kb, ein 0,9 kb, ein 0,8 kb, ein 0,65 kb und ein 0,5 kb langes SacI-Subfragment sowie ein 6,5 kb und ein 4,3 kb langes XhoI-Subfragment erhalten.

Die Subklonierung des ca. 13,2 kb großen Inserts des Phagenklons P5 erfolgte durch Verdau mit den Restriktionsenzymen SacI bzw. XhoI. Insgesamt wurden SacI-Fragmente von 6,5 kb, 3,3 kb, 3,2 kb, 0,8 kb und 0,3 kb Größe sowie XhoI-Fragmente von 7 kb und 3,2 kb Größe subkloniert.

5

Zur Klonierung des 3' von Phagenklon P17 und 5' von Phagenklon P2 gelegenen hTC-genomischen Sequenzbereichs wurden 3 Genomic Walkings mit Hilfe des GenomeWalker™ Kits der Firma Clontech (Katalognummer K1803-1) und verschiedenen Primerkombinationen durchgeführt. In einem Endvolumen von 50 µl wurde 1 µl humaner GenomeWalker Library HDL (Fa. Clontech) mit 10 pmol dNTP-Mix versetzt und in 1xKlen Taq PCR-Reaktionspuffer und 1xAdvantage Klen Taq Polymerase Mix (Fa. Clontech) eine PCR-Reaktion durchgeführt. Als Primer wurden 10 pmol eines internen genspezifischen Primers sowie 10 pmol des Adaptor Primers AP1 (5'-GTAATACGACTCACTATAGGGC-3'; Fa. Clontech) zugefügt.

10 Die PCR wurde als Touchdown-PCR in 3 Schritten durchgeführt. Zunächst wurde über 7 Zyklen für 20 sec bei 94°C denaturiert und anschließend für 4 min bei 72°C die Primer angelagert und die DNA-Kette verlängert. Es folgten 37 Zyklen bei denen für 20 sec die DNA bei 94°C denaturiert wurde, die anschließende Primerverlängerung aber für 4 min bei 67°C erfolgte. Abschließend folgte eine Kettenverlängerung für 4 min bei 67°C. Im Anschluß an diese erste PCR wurde das PCR-Produkt 1:50 verdünnt. Ein µl dieser Verdünnung wurde in einer zweiten „nested“ PCR zusammen mit 10 pmol dNTP-Mix in 1xKlen Taq PCR-Reaktionspuffer und 1xAdvantage Klen Taq Polymerase-Mix sowie 10 pmol eines „nested“ genspezifischen Primers und 10 pmol des „nested“ Marathon Adaptor Primers AP2 (5'-ACTATAGGGCACGCGTGGT-3'; Fa. Clontech) eingesetzt. Die PCR-Bedingungen entsprachen den in der ersten PCR gewählten Parametern. Als einzige Ausnahme wurden im ersten PCR-Schritt statt 7 Zyklen nur 5 Zyklen gewählt und im zweiten PCR-Schritt statt 37 Zyklen nur 24 Zyklen durchlaufen. Produkte dieser Nested-GenomicWalking-PCR wurden in den TA-Cloning Vektor pCRII der Fa. InVitrogen

15 20 25 30 kloniert.

Im ersten Genomic Walking wurde der genspezifische Primer C3K2-GSP1 (5'-GACGTGGCTCTTGAAGGCCTTG-3') sowie der „nested“ genspezifische Primer C3K2-GSP2 (5'-GCCTTCTGGACCACGGCATAACC-3') zusammen mit der HDL-Library 4 eingesetzt und ein 1639 bp langes PCR-Fragment erhalten. Im zweiten
5 Genomic Walking wurde mit dem genspezifischen Primer C3F2 (5'-CGTAGTTGAGCACGCTGAACAGTG-3') und dem „nested“ genspezifischen Primer C3F (5'-CCTTCACCCTCGAGGTGAGACGCT-3') aus der HDL-Library 4 ein PCR-Fragment von 685 bp Länge amplifiziert. Der dritte Genomic Walking
10 Ansatz führte unter Einsatz des genspezifischen Primers DEL5-GSP1 (5'-GGTGGATGTGACGGGCGCGTACG-3') und des „nested“ genspezifischen Primers C5K-GSP1 (5'-GGTATGCCGTGGTCCAGAAGGC-3') zur Klonierung eines 924 bp PCR-Fragments aus der HDL-Library 1. Insgesamt wurden durch dieses Genomic Walking-Verfahren 2100 bp der 3' von Phagenklon P17 gelegenen genomischen hTC-Region identifiziert (s. Fig. 7).

15 Die subklonierten Fragmente sowie die Genomic Walking-Produkte wurden einzelsträngig sequenziert. Unter Verwendung der Lasergene Biocomputing Software (DNASTAR Inc. Madison, Wisconsin, USA) wurden überlappende Bereiche identifiziert und Contigs gebildet. Insgesamt wurden aus den gesammelten
20 Sequenzen der Phagenklone P12, P17, P2, P3 und P5 sowie den Sequenzdaten aus dem Genomic Walking 2 große Contigs zusammengestellt. Contig 1 besteht aus Sequenzdaten von Phagenklon P12, P17 und den Sequenzdaten aus dem Genomic Walking. Contig 2 wurde aus den Sequenzen von Phagenklon P2, P3 und P5 zusammengesetzt. Überlappende Phagenklonbereiche sind in Fig. 7 schematisch
25 dargestellt. Die Sequenzdaten der 2 Contigs sind nachfolgend dargestellt. Das ATG Startcodon in Contig 1 ist unterstrichen. Das TGA Stopcodon ist in Contig 2 unterstrichen.

Contig1:

	ACTTGAGCCC	AAGAGTTCAA	GGCTACGGTG	AGCCATGATT	GCAACACCAC	ACGCCAGCCT	TGGTGACAGA	70
	ATGAGACCCT	GTCTCAAAAA	AAAAAAGGAA	AATTGAAATA	ATATAAGACA	TCTTCTCTGG	CCACAGTGGG	140
5	ACAAAACAGG	AAATCAACAA	CAAGAGGAAT	TTTGAAAAC	ATACAAACAC	ATGAAAATTA	AACAATATAC	210
	TTCTGAATGA	CCAGTGAGTC	AATGAAGAAA	TTAAAAAGGA	AATTGAAAAA	TTTATTTAAG	CAATGTATAA	280
	CGGAAACATA	ACCTCTCAAA	ACCCACGGTA	TACAGCAAAA	GCAGTGCTAA	GAAGGAAAGT	TATAGCTATA	350
	AGCAGCTACA	TCAAAAAAGT	AGAAAAAGCA	GGCGCAGTGG	CTCATGCGCTG	TAATCCCAGC	ACTTTGGGAG	420
	GCCAAGGCGG	GCAGATCGCC	TGAGGTCAAG	AGTTTCAGAC	CAGCCTGACC	AACACAGAGA	AACCTTGTCC	490
10	CTACTAAAAA	TACAAAATTA	GCTGGGCATG	GTGGCACATG	CCTGTAATCC	CAGCTACTCG	GGAGGCTGAG	560
	GCAGGATAAC	CGCTTGAACC	CAGGAGGTGG	AGGTTGCGGT	GAGCCGGGAT	TGCGCCATTG	GACTCCAGCC	630
	TGGGTAAACA	GAGTGAAACC	CTGTCTCAAG	AAAAAAGGAA	AAGTAGAAAA	ACTTAAAAAT	ACAACTTAAT	700
	GATGCACCTT	AAAGAACTAG	AAAAGCAAGA	GCAAACTAAA	CCTAAAAATTG	GTAAGAGAAA	AGAAATAATA	770
	AAGATCAGAG	CAGAAATAAA	TGAAACTGAA	AGATAACAAT	ACAAAAGATC	AACAAAATTA	AAAGTTGGTT	840
15	TTTTGAAAAG	ATAAACAAAA	TTGACAAACC	TTTGCCAGAA	CTAAGAAAAA	AGGAAAGAA	ACCTAAATAA	910
	ATAAAGTCAG	AGATGAAAAA	AGAGACATTA	CAACTGATAC	CACAGAAATT	CAAAGGATCA	CTAGAGGCTA	980
	CTATGAGCAA	CTGTACACTA	ATAAATGAAA	AAACCTAGAA	AAAATAGATA	AATTCCTAGA	TGCATACAAC	1050
	CTACCAAGAT	TGAACCATGA	AGAAATCCAA	AGGCCAAACA	GACCAATAAC	AATAATGGGA	TATAAGCCAT	1120
20	AATAAAAAAGT	CTCTTAGCAA	AGAGAAGCCC	AGGACCCAAT	GGCTTCCCTG	CTGGATTTTA	CCAATCATTT	1190
	AAAGAAGAAT	GAATTTCCAAT	CCTACTCAAA	CTATTCTGAA	AAATAGAGGA	AAGAATACTT	CCAAACTCAT	1260
	TCTCATGGCC	CAGTATTACC	CTGATTCCAA	AACCAGACAA	AAACACATCA	AAAACAAACA	AACAAAAAAA	1330
	CAGAAAGAAA	GAAACTTACA	GGCCAAATATC	CCTGATGAAT	ACTGATACAA	AAATCCTCAA	CAAAACACTA	1400
	GCAAAACCAA	TTAAACAACA	CAATTCGAAA	ATCATTCAAT	GTGATCAAGT	GGGATTTATT	CCAGGGATGG	1470
25	AAGGATGGTT	CAACATATGC	AAATCAATCA	ATGTGATACA	TCATCCCAAC	AAAATGAAGT	CAAAAAACTA	1540
	TATGATTATT	TCACTTTATG	CAGAAAAAGC	ATTTGATAAA	ATTCTGCACC	CTTCATGATA	AAAAACCTCA	1610
	AAAAACCAAG	TATACAAGAA	ACATACAGGC	CAGGCACAGT	GGCTCACACC	TGCGATCCCA	GCACCTCTGG	1680
	AGGCCAAGGT	GGGATGATTG	CTTGGGCCCA	GGAGTTTGAG	ACTAGCCTGG	GCAACAAAAT	GAGACCTGGT	1750
	CTACAAAAAA	CTTTTTTAAA	AAATTAGCCA	GGCATGATGG	CATATGCCCTG	TAGTCCAGC	TAGTCTGGAG	1820
30	GCTGAGGTGG	GAGAATCACT	TAAGCCTAGG	AGGTCGAGGC	TGCAGTGAGC	CATGAACATG	TCAGTGACT	1890
	CCAGCCTAGA	CAACAGAACCA	AGACCCCACT	GAATAAGAA	AAGGAGAAGG	AGAAAGGAGA	AGGGAGGGAG	1960
	AAGGAGGAGG	GAGGAGAAGG	AGGAGGTGGA	GGAGAAAGTG	AAGGGGAAGG	GGAAAGGAAA	AGGAAAGAA	2030
	AAGAAACATA	TTTCAACATA	ATAAAGCCCC	TATATGACAG	ACCGAGGTAG	TATTATGAGG	AAAAACTGAA	2100
	AGCCTTTCCCT	CTAAGATCTG	GAAATGACA	AGGGCCCACT	TTCAACCACTG	TGATTCAACA	TAGTACTAGA	2170
	AGTCCTAGCT	AGAGCAATCA	GATAAGAGAA	AGAAATAAAA	GGCATCCAAA	CTGGAAGGA	AGAAGTCAAA	2240
35	TTATCCTGTT	TGCAGATGAT	ATGATCTTAT	ATCTGGAATA	GACTTAAGAC	ACCACTAAAA	AACATTATGA	2310
	GCTGAAATTT	GGTACAGCAG	GATACAAAA	CAATGTACAA	AAATCAGTAG	TATTTCTATA	TTCCAACAGC	2380
	AAACAATCTG	AAAAAGAAAC	CAAAAAAGCA	GCTACAAATA	AAATTAACAA	GCTAGGAATT	AACCAAGAAA	2450
	GTGAAAGATC	TCTACAATGA	AAACTATAAA	ATGTTGATAA	AAGAAATTGA	AGAGGGCACA	AAAAAAGAAA	2520
40	AGATATTCCA	TGTTTCATAGA	TTGGAAGAA	AAATCTGTT	AAATGTGCTA	TACTACCCAA	AGCAATTTC	2590
	AAATTCATATG	CAATCCCTAT	TAAATACTA	ATGACGTTCT	TCACAGAAAT	AGAAAGAAACA	ATTCTAAGAT	2660
	TTGTACAGAA	CCACAAAAGA	CCCAGAATAG	CCAAAGCTAT	CCTGACCAAA	AAGAACAAA	CTGGAAGCAT	2730
	CACATTACCT	GACTTCAAT	TATACTACAA	AGCTATAGTA	ACCAAACTA	CATGGTACTG	GCATAAAAAC	2800
	AGATGAGACA	TGGACCAGAG	GAACAGAATA	GAGAAATCCG	AAACAAATCC	ATGCATCTAC	AGTGAACTCA	2870
45	TTTTTGACAA	AGGTGCCAAG	AACATACTTT	GGGGAAAGAA	TAATCTCTTC	AATAAATGGT	GCTGGAGGAA	2940
	CTGGATATCC	ATATGCAAAA	TAACAATACT	AGAACTCTGT	CTCTCAGGAT	ATACAAAAGC	AAATCAAAAT	3010
	GGATGAAAGG	CTTAAATCTA	AAACCTCAAA	CTTTGCAACT	ACTAAAGAAA	AACACCGGAG	AACTCTCCCA	3080
	GGACATTGGA	GTGGGCAAG	ACTTCTTGAG	TAATTCCTCG	CAGGCACAGG	CAACCAAGG	AAACACAGAC	3150
	AAATGGGATC	ATATCAAGTT	AAAAAGCTTC	TGCCAGCAAA	AGGAAACAA	CAACAAAGAG	AAGAGACAAC	3220
50	CCACAGAATG	GGAGAATATA	TTTGCAAACT	ATTCATCTAA	CAAGGAATTA	ATAACCACTA	TATATAAGGA	3290
	GCTCAAACTA	CTCTATAAGA	AAAAACACCTA	ATAAGCTGAT	TTTCAAAAA	AAGCAAAAGA	TCTGGGTAGA	3360
	CATTTCCTCA	AATAAGTCAT	ACAAATGGCA	AACAGGCATC	TGAAAATGTG	CTCAACACCA	CTGATCATCA	3430
	GAGAAATGCA	AATCAAACT	ACTATGAGAG	ATCATCTCAT	CCCAGTTAAA	ATGGCTTTTA	TTCAAAAGAC	3500
	AGGCAATAAC	AAATGCCAGT	GAGGATGTGG	ATAAAGGAAA	ACCTTTGGAC	ACTGTTGGTG	GGAAATGAAA	3570
55	TTGCTACCAC	TATGGAGAAC	AGTTTGAAAG	TTCTCAAAA	AACTAAAAAT	AAAGCTACCA	TACAGCAATC	3640
	CCATTGCTAG	GTATATACTC	CAAAAAAGGG	AATCAGTGTA	TCAACAAGCT	ATCTCCACTC	CCACATTTAC	3710
	TGCAGCACTG	TTCATAGCAG	CCAAGGTTTG	GAAGCAACCT	CAGTGTCAT	CAACAGACGA	ATGGAAAAAG	3780
	AAAATGTGGT	GCACATACAC	AATGGAGTAC	TACGCAGCCA	TAAAAAAGAA	TGAGATCCTG	TCAGTTGCAA	3850
	CAGCATGGGG	GGCACTGGTC	AGTATGTTAA	GTGAAAATA	CCAGGCACAG	AAAGACAAAC	TTTTTATGTT	3920
	CTCCCTTACT	TGTGGGAGCA	AAAAATTAATA	CAATTGACAT	AGAAATAGAG	GAGAAATGGT	GTCTAGAGG	3990
60	GGTGGGGGAC	AGGGTGACTA	GAGTCAACAA	TAATTTATTG	TATGTTTAA	AATAACTAAA	AGAGTATAAT	4060
	TGGGTTGTTT	GTAACACAAA	GAAAGGATAA	ATGCTTGAAG	GTGACAGATA	CCCCATTTAC	CCTGATGTGA	4130
	TTATTACACA	TTGTATGCC	GTATCAAAAT	ATCTCATGTA	TGCTATAGAT	ATAAACCCCTA	CTATATTAAA	4200
	AATTAATAAT	TTAATGGCCA	GGCACGGTGG	CTCATGTCCG	TAATCCCAGC	ACTTTGGGAG	GGCCAGGCGG	4270
	GTGGATCAC	TGAGGTCAAG	AGTTTGAAAC	CAGTCTGGCC	ACCATGATGA	AACCTGTCT	CTACTAAAGA	4340
65	TACAAAAATT	AGCCAGGCGT	GGTGGCACAT	ACCTGTAGTC	CCAACTACTC	AGGAGGCTGA	GACAGGAGAA	4410
	TTGCTTGAAC	CTGGGAGGCG	GAGGTTCGAG	TGAGCCGAGA	TCATGCCACT	GCAGTGCAGC	CTGGGTGACA	4480
	GAGCAAGACT	CCATCTCAAA	ACAAAAACAA	AAAAAGAAAG	ATTAAAAATTG	TAATTTTAT	GTACCGTATA	4550
	AATATATACT	CTACTATATT	AGAAGTTAAA	AATTAACA	ATTATAAAG	GTAATTAACC	ACTTAATCTA	4620
70	AAATAAGAAC	AATGTATGTG	GGGTTCTAG	CTTCTGAAGA	AGTAAAGTT	ATGGCCACGA	AGCCAGAAAT	4690
	GTGAGGAGGG	AACAGTGGAA	GTTACTGTTG	TTAGACGCTC	ATACTCTCTG	TAAGTGACTT	AATTTTAACC	4760
	AAAGACAGGC	TGGGAGAAAT	TAAAGAGGCA	TTCTATAAGC	CCTAAAAACAA	CTGCTAATAA	GGTGAAAGG	4830
	TAATCTCTAT	TAATTTACCA	TAATTACAGA	TATCTCTAAA	ATCGAGCTGC	AGAATTGGCA	CGTCTGATCA	4900
	CACCGTCTCT	TCATTACCGG	TGCTTTTTTT	TTTGTGTGCT	TGGAGATTTT	CGATTGTGTG	TTCTGTGTTG	4970
75	GTTAAACTTA	ATCTGTATGA	ATCTGAAAC	GAAAAATGGT	GGTGATTTC	TCCAGAGAA	TTAGAGTACC	5040
	TGGCAGGAAG	CAGGTGGCTC	TGTGGACCTG	AGCCACTTCA	ATCTTCAAGG	GTCTCTGGCC	AAGACCCAGG	5110

	TGCAAGGCAG	AGGCCTGATG	ACCCGAGGAC	AGGAAAGCTC	GGATGGGAAG	GGGCGATGAG	AAGCCTGCCT	5180
	CGTTGGTGAG	CAGCGCATGA	AGTGCCCTTA	TTTACGCTTT	GCAAAGATTG	CTCTGGATAC	CATCTGGAAA	5250
	AGGCGGCCAG	CGGGAATGCA	AGGAGTCAGA	AGCCTCCTGC	TCAAACCCAG	GCCAGCAGCT	ATGGCGCCCA	5320
5	CCCGGGCGTG	TGCCAGAGGG	AGAGGAGTCA	AGGCACCTCG	AAGTATGGCT	TAAATCTTTT	TTTACCTGA	5390
	AGCAGTGACC	AAGGTGTATT	CTGAGGGAAG	CTTGAGTTAG	GTGCCTTCTT	TAAACAGAA	AGTCATGGAA	5460
	GCACCTCTCT	CAAGGGAAAA	CCAGACGCCC	GCTCTGCGGT	CATTACCTC	TTTCTCTCT	CCCTCTCTTG	5530
	CCCTCGCGGT	TTCTGATCGG	GACAGAGTGA	CCCCCGTGA	GCTTCTCCGA	GCCCGTGCTG	AGGACCTCT	5600
	TGCAAGGGG	TCCACAGACC	CCGCGCTGG	AGAGAGGAGT	CTGAGCCTGG	CTTAATAACA	AACCTGGATG	5670
10	TGGCTGGGGG	CGGACAGCGA	CGCGGGGATT	CAAAGACTTA	ATTCCATGAG	TAAATTCAAC	CTTTCCACAT	5740
	CCGAATGGAT	TTGGATTTTA	TCTTAATATT	TTCTTAAATT	TCATCAAATA	ACATTACAGGA	CTGCAGAAAT	5810
	CCAAAGGCGT	AAAACAGGAA	CTGAGCTATG	TTTGCCAAAG	TCCAAGGACT	TAATAACCAT	GTTCAGAGGG	5880
	ATTTTTCGCC	CTAAGTACTT	TTTATTGGTT	TTTCAAGGT	GGCTTAGGGT	GCAAGGGGAA	GTACACGAGG	5950
	AGAGGCTGCG	GCGGCAGGGC	TATGAGCAGG	GCAGGGCCAC	CGGGGAGAGA	GTCCCGGGCC	TGGGAGGCTG	6020
15	ACAGCAGGAC	CACTGACCGT	CCTCCTCGGG	AGCTGCCACA	TTGGGCAACG	GCAAGGGGGC	CACGCTGCGT	6090
	GTGACTCAGG	ACCCATACCC	GGCTTCTGCG	GCCCCCCAC	ACTAACCCAG	GAAGTCACGG	AGCTCTGAAC	6160
	CCGTGGAAAC	GAACATGACC	CTTGCTGCC	TGCTTCCCTG	GGTGGGTCAA	GGGTAAATGAA	TGGGTGTGCA	6230
	GGAAATGGCC	ATGTAAATTA	CACGACTCTG	CTGATGGGGA	CCGTTCCTTC	CATCTATTAT	CATCTTCACC	6300
	CCCAAGGACT	GAATGATTCC	AGCAACTTCT	TCGGGTGTGA	CAAGCCATGA	CAAAACTCAG	TACAAACACC	6370
20	ACTCTTTTAC	TAGGCCACAC	GAGCAGCGSC	CACACCCCTG	ATATATTAAG	AGTCCAGGAG	AGATGAGGCT	6440
	GCTTTTCAGC	ACCAGGCTGG	GGTGACAACA	CGCGCTGAAC	AGTCTGTTC	TCTAGACTAG	TAGACCTGG	6510
	CAGGCACTCC	CCCAGATTCT	AGGGCTCGT	TGCTGCTTCC	CGAGGGCGCC	ATCTGCCCTG	GAGACTCAGC	6580
	CTGGGTGCCC	ACACTGAGGC	CAGCCCTGTC	TCCACACCTT	CCGCTCCAG	GCCCTCAGCT	CTCCAGCAGC	6650
	TTCTTAAACC	CTGGGTGGGC	CGTGTCCAG	CGCTACTGTC	TCACCTGTTC	CACCTGTGCT	CTCTCAGCG	6720
25	ACGTAGCTCG	CACGGTTCC	CCTCACATGG	GGTGTCTGTC	TCCTTCCCCA	ACACTCACAT	CGGTGGAAGG	6790
	GAGGAGATTG	TGCGCCTCCC	AGACTGGGCT	CTCTGAGCCT	GAACCTGGCT	CGTGGCCCCC	GATCAGGCT	6860
	CTTGGGCTCC	GGCTGCACGC	TGACCTCCAT	TTCCAGGCGC	TCCCGCTCTC	CTGTCTATCT	CCGGGGCCTG	6930
	CCGGTGTGTT	CTTCTGTTTC	TGTGCTCCTT	TCCACGTCCA	GCTGCGTGTG	TCTCTGCCCG	CTAGGGTCTC	7000
	GGGGTTTTTA	TAGGCATAGG	ACGGGGGCGT	GGTGGGCCAG	GGCGCTCTTG	GGAAATGCAA	CATTGGGGTG	7070
30	TGAAAGTAGG	AGTGCCGTGC	CTCACCTAGG	TCCACGGGCA	CAGGCCTGGG	GATGGAGCCC	CCGCCAGGGA	7140
	CCCGCCCTCT	TCTGCCACGC	ACTTTCCTCG	CCCCCTCCCT	CTGGAAACACA	GAGTGGCAGT	TTCCACAAGC	7210
	ACTAAGCATC	CTCTTCCCAA	AAGACCCAGC	ATTGGCACCC	CTGGACATTT	GCCCCACAGC	CCTGGGAATT	7280
	CAGCTGACTA	CGCACATCAT	GTACACACTC	CCGTCCACGA	CCGACCCCGC	CTGTTTTATT	TAAATAGCTA	7350
	CAAGAGCAGG	AAATCCCTGC	TAAATGTCC	TTTAAACAAAC	TGGTTAAACA	AACGGGTCCA	TCCGCACGGT	7420
35	GGACAGTTCC	TCACAGTGAA	GAGGAACATG	CCGTTTATAA	AGCCTGCAGG	CATCTCAAGG	GAATTACGCT	7490
	GAGTCAAAAC	TGCCACCTCC	ATGGGATACG	TACGCAACAT	GCTCAAAAAG	AAAGAAATTC	ACCCATGGC	7560
	AGGGGAGTGG	TTAGGGGGGT	TAAAGACGGT	GGGGGCGGCA	GCTGGGGGCT	ACTGCACGCA	CTTTTACTA	7630
	AAGCCAGTTT	CCTGGTTCTG	ATGGTATTGG	CTCAGTTATG	GGAGACTAAC	CATAGGGGAG	TGGGGATGGG	7700
40	GGAAACCCGA	GGCTGTGCCA	TCCTTGGCAT	GCCCGAGTGT	CCTGGGCAGG	ATAATGCTCT	AGAGATGCC	7770
	ACGTCCTGAT	TCCCCCAAAC	CTGTGGACAG	AACCCGCCCC	GCCCCAGGGC	CTTTGCAAGT	GTGATCTCCG	7840
	TAGGAGCCCT	GAGGTCTGGG	ATCCTCTGGG	ACTACCTGCA	GGCCGCAAAA	GTAATCCAGG	GGTCTGGGA	7910
	AGAGGCGGGC	AGGAGGGTCA	GAGGGGGGCA	GCCTCAGGAC	GATGGAGGCA	GTCAGTCTGA	GGCTGAAAAG	7980
	GGAGGGAGGG	CCTCGAGCCC	AGGCTTGCAA	GCGCCTCCAG	AAGCTGGAAA	AGCGGGGGAA	GGGACCTCC	8050
45	ACGGAGCCTG	CAGCAGGAAG	GCACGGCTGG	CCCTTAGCCC	ACCAGGGCCC	ATCGTGGACC	TCCGGCTCC	8120
	GTGCCATAGG	AGGGCACTCG	CGCTGCCCTT	CTAGCATGAA	GTGTGTGGGG	ATTTGCAGAA	GCAACAGGAA	8190
	ACCCATGAC	TGTGAATCTA	GGATTATTTT	AAACAAAGG	TTTACAGAAA	CATCCAAGGA	CAGGGCTGAA	8260
	GTGCCCTCCG	GCAAGGGCAG	GGCAGGCACG	AGTGATTTTA	TTTAGCTATT	TTATTTTATT	TACTTACTTT	8330
	CTGAGACAGA	GTTATGCTCT	TGTTGCCAG	GCTGGAGTGC	AGCGGCATGA	TCTTGGCTCA	CTGCAACCTC	8400
	CGTCTCCTGG	GTTCAAGCAA	TTCTCGTGCC	TCAGCCTCCC	AAGTAGCTGG	GATTTACAGC	GTGCACCACC	8470
50	ACACCCGGCT	AATTTGTGAT	TTTATAGTAG	GATGGGCTTT	CACCATGTTG	GTCAAGCTGA	TTCAAAATCT	8540
	CTGACCTCAG	GTGATCCGCC	CACCTCAGCC	TCCCAAAGTG	CTGGGATTAC	AGGCATGAGC	CACCTGCACCT	8610
	GGCCTATTTA	ACCATTTTAA	AACCTTCCCTG	GGCTCAAGTC	ACACCCACTG	GTAAGGAGTT	CATGGATGCT	8680
	AATTTCCCTT	TTACTCAGGA	GTTACCTCCC	TTTGATATTT	TCTGTAATTC	TTCTAGACT	GGGGATACAC	8750
	CGTCTCTTGA	CATATTACAC	GTTTCTGTGA	CCACCTGTTA	TCCCATGGGA	CCCAGTGCAG	GGGCAGCTGG	8820
55	GAGGCTGCAG	GCTTCAGGTC	CCAGTGGGGT	TGCCATCTGC	CAGTAGAAAC	CTGATGTAGA	ATCAGGGCGC	8890
	AAGTGTGGAC	ACTGTCTCTA	ATCTCAATGT	CTCAGTGTGT	GCTGAAACAT	GTAGAAATTA	AAGTCCATCC	8960
	CTCCTACTCT	ACTGGGATTG	AGCCCCCTCC	CTATCCCCCC	CCAGGGGCGAG	AGGAGTTCCCT	CTCACTCCTG	9030
	TGGAGGAAGG	AATGATACTT	TGTTATTTTT	CAGTGTGTGT	ACTGAATCCA	CTGTTTCATT	TGTTGGTTTG	9100
	TTTGTTTTGT	TTTGAGAGGC	GGTTTCACTC	TTGTTGCTCA	GGCTGGAGGG	AGTGCAATGG	CGCGATCTTG	9170
60	GCTTACTGCA	GCCTCTGCCT	CCCAGGTCCA	AGTGATTCTC	CTGCTTCCGC	CTCCCATTTG	GCTGGGATTA	9240
	CAGGCACCCG	CCACCATGCC	CAGCTAATTT	TTTGATTTTT	TAGTAGAGAC	GGGGGTGGGT	GGGGTTACCC	9310
	ATGTTGGCCA	GGCTGGTCTC	GAACCTCTGA	CCTCAGATGA	TCCACCTGCC	TCTGCCTCTC	AAAGTGTCTG	9380
	GATTACAGGT	GTGAGCCACC	ATGCCCCAGCT	CAGAATTTAC	TCTGTTTAGA	AACATCTGGG	TCTGAGGTAG	9450
	GAAGCTCACC	CCACTCAAGT	GTTGTGGTGT	TTTAAGCCAA	TGATAGAATT	TTTTTATTGT	TGTTAGAACA	9520
65	CTCTTGATGT	TTTACACTGT	GATGACTAAG	ACATCATCAG	CTTTTCAAAG	ACACACTAAC	TGCACCCATA	9590
	ATACTGGGGT	GTCTTCTGGG	TATCAGCAAT	CTTCATTGAA	TGCCGGGAGG	CGTTTCTCCG	CCATGCACAT	9660
	GGTGTAAATT	ACTCCAGCAT	AATCTTCTGC	TTCCATTCTT	TCTCTTCCCT	CTTTTAAAT	TGTTGTTTCT	9730
	ATGTTGGCTT	CTCTGCAGAG	AACCACTGTA	AGCTACAAC	TAACCTTTGT	TGGAACAAAT	TTTCCAAACC	9800
	GCCCCCTTGC	CCTAGTGGCA	GAGACAATTC	ACAAACACAG	CCCTTTAAAA	AGGCTTAGGG	ATCACTAAGG	9870
70	GGATTCTAG	AAGAGCGACC	TGTAATCTTA	AGTATTTACA	AGACGAGGCT	AACCTCCAGC	GAGCGTGACA	9940
	CCCCAGGGAG	GGTCCGAGGC	CTGTTCAAAT	GCTAGCTCCA	TAAATAAAGC	AATTTCTCTC	GGCAGTTTCT	10010
	GAAAGTAGGA	AAGGTTACAT	TTAAGGTTGC	GTTTGTTAGC	ATTTCAAGT	TTGCCGACCT	CAGCTACAGC	10080
	ATCCCTGCAA	GGCCTCGGGA	GACCCAGAAG	TTTCTCGCCC	CCTTAGATCC	AAACTTGAGC	AACCCGAGT	10150
	CTGGATTCTT	GGGAAGTCTT	CAGCTGTCTT	CGGTTGTGTC	CGGGGCCCA	GCTCTGGAGG	GGACAGTGG	10220
75	CCGTGTGGCT	TCTACTGCTG	GGCTGGAGT	CGGGCTCTCT	AGCTCTGACG	TCCGAGGCTT	GGAGCCAGT	10290
	GCCTTGAACCC	CGAGGCTGCC	CTGCCCTCTG	TGCGGGCGGG	ATGTGACACG	ATGTTGGCTC	CATCTGCCAG	10360
	ACAGAGTGCC	GGGGCCACAG	GTCAGGGCCG	TTGTGGCTGG	TGTGAGGCGC	CCGGTGGCGG	GCCAGCAGGA	10430
	GCGCTTGCT	CCATTTCCCA	CCCTTCTCTG	ACGGGACCCG	CCCGGTGGGT	GATTAACAGA	TTTGGGGTGG	10500

	TTTGCTCATG	GTGGGGACCC	CTCGCCGCC	GAGAACCTGC	AAAGAGAAAT	GACGGGCGCT	TGTCAAGGAG	10570
	CCCAAGTCGC	GGGGAAGTGT	TGCAGGGAGG	CACTCCGGGA	GGTCCCGCGT	GCCGCTCCAG	GGAGCAATGC	10640
	GTCCTCGGGT	TGCTCCCCAG	CCGCGTCTAC	GCGCTCCGT	CCTCCCTTC	ACGTCCGGCA	TTCGTGGTGC	10710
5	CCGGAGCCCG	ACGCCCCCGG	TCCGGACCTG	GAGGCAGCCC	TGGGTCTCCG	GATCAGGCCA	GCGGCCAAAG	10780
	GGTGCCCGCA	CGCACCTGTT	CCCAGGGCCT	CCACATCATG	GCCCCCTCC	CGGGTTACCC	CACAGCCTAG	10850
	GCCGATTCTGA	CCTCTCTCCG	CTGGGGCCCT	CGTGGCGCTC	CCTGCACCCT	GGGAGCGCGA	GCGGCGCGCG	10920
	GGCGGGGAAG	CGCGGGCCAG	ACCCCGGGGT	CCGCGCGGAG	CAGCTGCGCT	GTCGGGGCCA	GCGCCGGGCTC	10990
	CCAGTGGGATT	CGCGGGCACA	GACGCCCAGG	ACCGCGCTCC	CCACGTGGCG	GAGGAGACTG	GGACCCGGGC	11060
10	ACCCGTCTGT	CCCTTTCACC	TTCAGCTTCC	GCCTCTCCG	CGCGGACCCC	GCCCCGTCCC	GACCCCTCCC	11130
	GGGTCCCCGG	CCCAGCCCCC	TCCGGGCCCT	CCCAGCCCCC	CCCTTTCCTT	TCCCGGGCCC	CGCCCTCTCT	11200
	TCCGGGGCGG	AGTTTCAGCG	AGCGCTGCGT	CCTGCTGCGC	ACGTGGGAAG	CCCTTGGCCC	GGCCACCCCC	11270
	GCGATTGCGG	GCGCTCCCGG	CTGCCGAGCC	GTGCGCTCCC	TGCTGCGCAG	CCACTACCGC	GAGGTGCTGC	11340
	CGCTGGCCAC	GTTCTGCGGG	CGCCTGGGGC	CCCAGGGCTG	GCGGCTGGTG	CAGCGCGGGG	ACCCGGCGGC	11410
15	TTTCCGCGCG	CTGGTGGCCC	AGTGCTGGT	GTGCGTGCCC	TGGGACGCAC	GCGCGCCCCC	CGCCGCCCCC	11480
	TCCTTCCGCG	AGGTGGGCTT	CCCCGGGCTC	GGCGTCCGGC	TGGGTTGAG	GCGGCGCGGG	GGGAACCAGC	11550
	GACATGCGGA	GAGCAGCGCA	GGCGACTCAG	GGCGCTTCCC	CCCGAGGTGT	CCTGCTGAA	GGAGCTGGTG	11620
	GCCCGAGTGC	TGCAGAGGCT	GTGCGAGCGC	GGCGCGAAGA	ACGTGCTGGC	CTTCGGGCTC	GCGCTGCTGG	11690
	ACGGGGCCCG	CGGGGGCCCC	CCCAGGCCCT	TCACCAACAG	CGTGCGCAGC	TACCTGCCCA	ACACGGTGAC	11760
20	CGACGCACTG	CGGGGGAGCG	GGGCGTGGGG	GCTGCTGCTG	GCGCCGCTGG	GCGACGACGT	GCTGGTTTAC	11830
	CTGCTGGCAC	GCTGCGCGCT	CTTTGTGCTG	TGCGCTCCCA	GCTGCGCTTA	CCAGGTGTGC	GGGCGCGCGC	11900
	TGTACGAGCT	CGGCGCTGCC	ACTCAGGCCC	GGCCCCGCCC	ACACGCTAGT	GGACCCCGAA	GCGCTCTGGG	11970
	ATGCGAACGG	GCCTGGAACC	ATAGCCTCAG	GGAGGCCGGG	GTCCCCCTGG	GCCTGCCAGC	CCCCGGTGGC	12040
	AGGAGGCCGG	GGGGCAGTGC	CAGCCGAAGT	CTGCCGTGGC	CCAAGAGGCC	CAGGCGTGGC	GCTGCCCTGT	12110
25	AGCCGGAGCG	GACGCCCGTT	GGGCGAGGGT	CCTGGGCCCC	CCCGGGCAGG	ACCGGTGGAC	CGAGTGACCG	12180
	TGGTTTCTGT	TGGGTGTAC	CTGCCAGACC	CGCGGAAGAA	GCCACCTCTT	TGGAGGGTGC	GCTCTGTGGC	12250
	ACGCGCCACT	CCCACCCATC	CGTGGGCGGC	CAGCACCAGC	CAGGCGCCCC	ATCCACATCG	GGGCCACCAC	12320
	GTCCCTGGGA	CACGCCCTGT	CCCCGGGTGT	ACGCCGAGAC	CAAGCACTTC	CTCTACTCCT	CAGGCGACAA	12390
	GGAGCAGCTG	CGGCCCTCCT	TCCTACTCAG	CTCTCTGAGG	CCCAGCCTGA	CTGGCGCTCG	GAGGCTCTGT	12460
30	GAGACCATCT	TTCTGGGTTT	CAGGCCCTGG	ATGCCAGGGA	CTCCCGCAG	GTTGCCCGCG	CTGCCCGTCT	12530
	GCTACTGGCA	AATGCGGCCC	CTGTTTCTGG	AGCTGCTTGG	GAACACGCGC	CAGTGCCCTT	ACGGGGTGCT	12600
	CCTCAAGACG	CACTGCCCGG	TGCGAGTCTC	GGTCACCCCA	GCAGCCGGTG	TCTGTGCCCG	GGAGAAGGCC	12670
	CAGGGCTCTG	TGGCGGCCCG	CAGGAGGAGG	GACACAGACC	CCCGTGCCTT	GSTGAGCTG	CTCCGCCAGC	12740
	ACAGCAGCCC	CTGGCAGGTG	TACGGCTTCG	TGCGGGCGCT	CCTGCGCGCG	CTGGTGCCCC	CAGGCTCTGT	12810
35	GGGCTCCAGG	CACAACGAAC	GCCGCTTCTT	CAGGAACACC	AAGAAGTTCA	TCTCCCTGGG	GAAGCATGCC	12880
	AAGCTCTCGC	TGCAGGAGCT	GACGTGGAAG	ATGAGCGTGC	GGGACTGCGC	TTGCGCTGCG	AGGAGCCGAG	12950
	GTGAGGAGGT	GGTGGGCGTG	GAGGGGCCAG	GCCCCAGAGC	TGAATGCAGT	AGGGGCTCAG	AAAAGGGGGC	13020
	AGGCAGAGCC	CTGGTCTCCT	GTCACGTGGG	CACACGTGGC	TTTTCGCTCA	GGACGTGCTG	CAGGAGCTGC	13090
	TGGACACGGT	GATCTCTGCC	TCTGCTCTCC	CTCCTGTCCA	GTTTGCATAA	ACTTACGAGG	TTCACTTCA	13160
40	CGTTTTGATG	GACACCGCGT	TTCAGGCGCC	CGAGGCCAGA	GCAGTGAACA	GAGGAGGCTG	GGCGGGCAG	13230
	TGGAGCCGGG	TTGCCGGCAA	TGGGGAGAAG	TGCTGGAAG	CACAGACGCT	CTGCCGAGGG	TGCTTGCAGG	13300
	TTACCTATAA	TCCTCTTCGC	AATTTCAAGG	GTGGGAATGA	GAGGTGGGGA	CGAGAACCCC	CTCTTCTCTG	13370
	GGGTGGGAGG	TAAAGGTTTT	GCAGGTGCAC	GTGGTCAGCC	AATATGCAGG	TTTGTGTTTA	AGATTTAATT	13440
	GTGTGTGAC	GGCCAGGTGC	GGTGGCTCAC	GCCGGTAATC	CCAGCACTTT	GGGAAGCTGA	GGCAGGTGGA	13510
45	TCACCTGAGG	TACAGAGTTT	GAGACCAAGC	TGACCAACAT	GGTGAACCCC	TATCTGTACT	AAAAATACAA	13580
	AAATTAGCTG	GGCATGGTGG	TGTGTGCGCT	TAATCCCAGC	TACTTGGGAG	GCTGAGGCAG	GAGATCACT	13650
	TGAACCCAGG	AGGCGGAGGC	TGAGATTGTG	CCATTGTACT	CCAGCCTGGG	GCACAAGAGT	13720	
	GAAACTCTGT	CTTTAAAAAA	AAAAAGTGT	CGTTGATTGT	GCCAGGCAG	GSTAGAGGGA	GGGAGATAAG	13790
	ACTGTTCTCC	AGCACAGATC	CTGGTCCCAT	CTTTAGGTAT	GAAGAGGGCC	ACATGGGAGC	AGAGGACAGC	13860
50	AGATGGCTCC	ACCTGCTGAG	GAAGGGACAG	TGTTTGTGGG	TGTTCAAGGG	ATGGTGCTGC	TGGGCCCTGC	13930
	CGTGTCGCCA	CCCTGTTTTT	CTGGATTGGA	TGTTGAGGAA	CCTCCGCTCC	AGCCCCCTTT	TGGTCCCGAG	14000
	TGCTCCCAAG	CCCTACCGTG	GCAGCTAGAA	GAAGTCCCGA	TTTCAACCCC	TCCCCACAAA	CTCCCAAGAC	14070
	ATGTAAGACT	TCCGGCCATG	CAGACAAGGA	GGGTGACCTT	CTTGGGGCTC	TTTTTTTTCT	TTTTTTCTTT	14140
	TTATGGTGGC	AAAAGTCATA	TAACATGAGA	TGGCACTCC	TAACACCGTT	TTCTGTGTAC	AGTGCAGAA	14210
55	TGCTAACTCG	CGCGTGTGTA	CAGCAGGTG	CTTGAAATGC	TGCGTCTTGC	GTGACTGGAA	GTCCCTACCC	14280
	ATCGAACGGC	AGCTGCCTCA	CACCTGCTGC	GGCTCAGGTG	GACCACGCCG	AGTCAGATAA	CGCTCATGCA	14350
	ACCCAGTTTT	GCTTTTTGTG	CTCCAGCTTC	CTTCGTTGAG	GAGAGTTTGA	GTTCTCTGAT	CAGGACTCTG	14420
	CCTGTCAATTG	CTGTTCTCTG	ACTTCAGATG	AGGTCACAAT	CTGCCCTTGG	CTTATGCAGG	GAGTGAGGCG	14490
	TGGTCCCCGG	GTGTCCCTGT	CACGTGCAGG	GTGAGTGAGG	CGTTGCCCCC	AGGTGTCCCT	GTACGCTGTA	14560
60	GGGTGAGTGA	GGCGCGGCCC	CCGGGTGTCC	CTGTCCCGTG	CAGCGTGATT	GAGGTGTGGC	CCCCGGGTGT	14630
	CCCTGTACCG	TGTAGGGTGA	GTGAGCGGCC	ATCCCCGGGT	GTCCCTGTCA	CGTGTAGGCT	GAGTGAGGCG	14700
	TGGTCCCCGG	GTGTCCCTGT	CCCGTGCAGG	GTGAGTGAGG	CAGTGTCCCC	GGGTGTCCCT	GTACGCTGCA	14770
	GGGTGAGTGA	GGCGCGGTCC	CCGGGTGTCC	CTCTCAGGTG	TAGGGTGAGT	GAGGCGCGGC	CCCAGGGTGT	14840
	CCCTGTACCG	TGTAGGGTGA	GTGAGGCACC	GTCCCTGGGT	GTCCCTCCCA	GGTATAGGGT	GAGTGAGGCA	14910
65	CTGTCCCCGG	GTGTCCCTGT	CACGTGCAGG	GTGAGTGAGG	CGCGGCCCCC	GGGTGTCCCT	CTCAGGTGCA	14980
	GGGTGAGTGA	GGCGCTGTCC	CTGGGTGTCC	CTGTCTCGTG	TAGGGTGAGT	GAGGCTCTGT	CCCCAGGTGT	15050
	CCTTGGCGTT	TGCTCACTTG	AGCTTGCTCC	TGAATGTTTG	CTCTTTCTAT	AGCCACAGCT	GCGCGGGTGT	15120
	CCCAATTGCT	GGGTAGATGG	TGCAGGCACA	GTGCTGGTCC	CCAAGCTAT	CTTTCTGTAT	GCTCGGCTCT	15190
	TCTTGGTCAC	CTCTCCGTTT	CATTTTGCTA	CGGGGACACG	GGACTGCAGG	CTCTCGCTCC	CGCGTGCCCA	15260
70	GGCACTGCAG	CCACAGCTTC	AGGTCCGCTT	GCCTCTGTTG	GGCCTGGCTT	GCTCACCAGC	TGCCCGCCAC	15330
	ATGCATGCTG	CCAATACTCC	TCTCCAGCT	TGTCTCATGC	CGAGGCTGGA	CTCTGGGCTG	CCTGTGTCTG	15400
	CTGCCACGTG	TGCTGGGAGA	CATCCAGAA	AGGGTTCTCT	GTGCCCTGAA	GGAAAGCAAG	TCACCCACGC	15470
	CCCCCACTTT	GTCTGTGTTT	CTCCCAAGCT	GCCTCTGTC	TTGGCCCCCT	TGGGTGGGTG	GCAACGCTTG	15540
	TCACCTTATT	CTGGGCACCT	GCCGCTCATT	GCTTAGGCTG	GGCTCTGCGT	CCAGTCGCCC	CCTCACATGG	15610
	ATTGACGTCC	AGCCACAGGT	TGGAGTGTCT	CTGTCTGTCT	CCTGCTCTGA	GACCCACGTG	GAGGGCCGGT	15680
75	GTCTCCGCCA	GCCTTCGTCA	GACTTCCCTC	TGGGCTTTTA	GTTTGAATTT	TCAGTGATT	ACCTCTGACG	15750
	TTTCTATCTC	TCCATTGTAT	GCTTTTTCTT	GGTTTATTCT	TTCAATTCCT	TTCTAGCTTC	TTAGTTTATG	15820
	CATGCCCTTC	CCTCTAAGTG	CTGCCCTTACC	TGCACCCCTG	GTTTTGATGT	GAAGTAATCT	CAACATCAGC	15890

	CACCTTCAAG	TGTTCTTAAA	ATACTTCAAA	GTGTTAATAC	TTCTTTTAAAG	TATCTTTATT	CTGTGATTTT	15960
	TTTCTTTGTC	CACGCTGTGT	TTTGACGTGA	AATCATTTTG	ATATCAGTGA	CTTTTAAAGTA	TTCTTTAGCT	16030
	TATCTGTGTA	TTTCTTTGAG	CAGTGAGTTA	TTTGAACACT	GTTTATGTTC	AAGATATGTA	GAGTATCAAG	16100
	ATACGTAGAG	TATTTTAAAGT	TATCATTTTA	TTATTGATTT	CTAACTCAGT	TGTGTAGTGG	TCTGTATAAT	16170
5	ACCAATTATT	TGAAGTTTGC	GGAGCCTTGC	TTTGTGATCT	AGTGTGTGCA	TGGTTTCCAG	AACGTGTCAT	16240
	TGTTAAATTTG	ACATCCTGTC	AATAGTGGGC	ATGCATGTTT	ACTATATCCA	GCTTATTAAG	GTCCAGTGCA	16310
	AAGCTTCTGT	CTCCTTCTAG	ATGCATGAAA	TTCCAAGAAG	GAGGCCATAG	TCCCTCACCT	GGGGGATGGG	16380
	TCTGTTCATT	TCTTCTCGTT	TGGTAGCATT	TATGTGAGGC	ATTGTTAGGT	GCATGCACGT	GGTAGAATTT	16450
10	TTATCTTCCT	GATGAGTGAA	TCTTTGGAG	ACTTCTATGT	CTCTAGTAAT	CTAGTAATTC	TTTTTTTAAA	16520
	TGTGCTTTAG	TACTGCCACA	CTGGGCTTCT	TTTGATTAGT	ATTTTCTGTC	TGTGCTGTTT	TTCTGCCTTT	16590
	AAATTATATA	TATATATATA	TTTTTTTTTT	TTTTGAGACA	GAGTCTTGGT	CTGTCCGCCA	GGGTGAGTGC	16660
	AGTGGTGTGA	TCACAGGTCA	GTGTAACCTT	TACCTTCTGG	CCTGAGCCGT	CCTCTCACCT	CAGCCTCCTG	16730
	AGTAGCTGGA	ACTGCAGACA	CGCACCGCTA	CACCTGGCTA	ATTTTAAAT	TTTTTCTGGA	GACAGGGTCT	16800
15	TGCTGTGTTG	CCCAGGCTGG	TCTCAAACTC	TTGGACTCAA	GGGATCCATC	TACCTCGGCT	TCCCAAAGTG	16870
	CTGAATTACA	GGCATGAGCC	ACCATGTCTG	GCCTAAATTT	CAACACTTTT	ATATCTTAT	AGTGTTGGTA	16940
	TGCTCTGTTA	ACAGCATGTA	GGTGAAATTC	CAATCCAGTC	TGACAGTCGT	TGTTTAACTG	GATAACCTGA	17010
	TTTATTTTCA	TTTTTTTGTG	ACTAGAGACC	CGCCTGGTGC	ACTCTGATTC	TCCACTTGCC	TGTTGATGTT	17080
	CCTCGTTCCC	TTGTTTCTCA	CCACCTCTTG	GGTTGCCATG	TGCGTTTCTT	GCCGAGTGTG	TGTTGATCCT	17150
20	CTCGTTGCC	CCTGGTCACT	GGGCAATTGC	TTTTATTCTT	CTTTGCTTAG	TGTTAGCCCC	TGATCTTTTT	17220
	ATTGTGCTTG	TTTGCTTTTG	TTTATTGAGA	CAGTCTCACT	CTGTCAACCA	GGCTGGAGTG	TAATGGCACA	17290
	ATCTCGGCTC	ACTGCAACCT	CTGCCCTCTC	GGTTCAAGCA	GTTCCTCATC	CTCAACCTCA	TGAGTAGCTG	17360
	GGATTACAGG	CGCCACACAC	CACGCTGGGC	TAATTTTTGT	ATTTTATGTA	GAGATAGGCT	TTCACCATGT	17430
	TGGCCAGGCT	GGTCTCAAA	TCTGACCTC	AAGTGATCTG	CCCGCTTGGG	CCTCCACAG	TGCTGGGATT	17500
25	ACAGGTGCAA	GCCACCGTGC	CCGGCATACC	TTGATCTTTT	AAAAAGAAAGT	CTGAAACATT	GCTACCCCTG	17570
	TCCTGAGCAA	TAAGACCCCT	AGTGATTTTT	AGCTCTGGCC	ACCCCCAGC	CTGTGTGCTG	TTTTCCCTGC	17640
	TGACTTAGTT	CTATCTCAGG	CATCTTGACA	CCCCACAAG	CTAAGCATT	TAAATATTGT	TTTCCGTGTT	17710
	GAGTGTTCCT	GTAGCTTTGC	CCCCGCCCTG	CTTTTCTCTC	TTTGTTCCTC	GTCTGTCTTC	TGCTCAGGCT	17780
	CGCCGCTCTG	GGGTCCCTTT	CTTGTCTCTT	TGCGTGGTTC	TTCTGTCTTG	TTATTGCTGG	TAAACCCAG	17850
30	CTTTACCTGT	GCTGGCTTCC	ATGGCATCTA	GCGACGCTCG	GGGACCTCTG	CTTATGATCG	ACAGATGAAG	17920
	ATGTGGAGAG	TCACAGGAGG	GGCGGTCATC	TTGGCCGCTG	AGTGTCTGGA	GCACACAGTG	GCCAGGCTTC	17990
	CTTAGCCAGT	GAGTGACAGC	AACGTCGCT	CGGCTGGGT	TCAGCCTGGA	AAACCCAGG	CATGTCGGGG	18060
	TCTGGTGGCT	CGCGGTGTGC	GAGTTTGAAA	TCGCGCAAA	CTGCGGTGTG	GCGCCAGCTC	TGACGGTGTCT	18130
	GCCTGGCGGG	GGAGTGTCTG	CTTCTCTCCT	TCTGCTTGGG	AACCAGGACA	AAGGATGAGG	CTCCGAGCCG	18200
35	TTGTGCGCCA	ACAGGAGCAT	GACGTGAGCC	ATGTGGATAA	TTTTAAATTT	TCTAGGCTGG	GCGCGGTGGC	18270
	TCACGCCCTGT	AATCCAGCA	CTTTGGGAGG	CCAAGGCGGG	TGGATCAGCA	GGTCAGGAGG	TCGAGACCAT	18340
	CCTGGCCAA	ATGATGAAC	CCCATCTGTA	CTAATAACAC	AAAAATTAGC	TGGGCGTGGT	GGCGGGTGCC	18410
	TGTAATCCCA	GCTACTCGGG	AGGCTGAGGC	AGGAGAATTG	CTTGAACCTG	GGAGTTGGAA	GTTGACGTGA	18480
	GCCGACATTG	CACCACTGCA	CTCCAGCCTG	GCAACACAGC	GAGACTCTGT	CTCAAAAAAA	AAAAAAAAAA	18550
40	AAAAAATAAA	AATTTAGTA	GCCACATTAA	AAAAAGTAAA	AAGAAAAGGT	GAAATTAATG	TAATAATAGA	18620
	TTTTACTGAA	GCCAGCATG	TCCACACCTC	ATCATTTTAG	GGTGTATTG	GTGGAGCAT	CACCTCAGAG	18690
	ACATTTGACA	TTTTTTGAGC	TTTGTCTGCG	GGATCCCGTG	TGTAGGTCCC	GTGCGTGGCC	ATCTCGGCCCT	18760
	GGACCTGCTG	GGCTTCCCAT	GGCCATGGCT	GTTGTACACG	ATGGTGACAG	TCCGGGATGA	GGTCGCCAGG	18830
	CCCTCAGTGA	GCTGGATGTG	CAGTGTCCGG	ATGGTGACAG	TCTGGGATGA	GGTCGCCAGG	CCCTGCTGTT	18900
45	AGCTGGATGT	GTGGTGTCTG	GATGGTGACG	GTCAGGGGTG	AGGTCTCCAG	GCCCTCGGTG	AGCTGGAGGT	18970
	ATGGAGTCCG	GATGATGCAG	GTCCGGGGTG	AGGTCTCCAG	GCCCTGCTGT	GAGCTGGATG	TGTGGTGTCT	19040
	GGATGGTGCA	GGTCAGGGGT	GAGGTCTCCA	GGCCCTCGGT	AAGCTGGAGG	TATGGAGTCC	GGATGATGGA	19110
	GGTCCGGGGT	GAGGTGCGCA	GGCCCTGCTG	TGAGCTGGAT	GTGTGGTGTG	TGGATGGTGC	AGGCTCTGGG	19180
	TGAGGTCAAC	AGGCCCTGCG	GTGAGCTGGG	TGTGCGGTGT	CTGGATGGTG	CAGGTCTGGA	GTGAGGTCCG	19250
50	CAGACGGTGC	CAGACCATGC	GGTGAGCTGG	ATATGCGGTG	TCCGGATGGT	GCAGGTCTGG	GGTGAGGTGT	19320
	CCAGGCCCTG	CTGTGAGTTG	GATGTGGGGT	GTCCGGATGC	TGCAGGTCCG	GTGTGAGGTG	ACCAGGCCCT	19390
	GCTGTGAGCT	GGATGTGTGG	TGTCTGGATG	GTGCAAGTCT	GGGGTGAAGG	TCCGACGGCC	CTGCTTGTG	19460
	AGCTGGATGT	GTGGTGTCTG	GATGGTGACG	GTCTGGAGTG	AGGTCTCCAG	GCCCTCGGTG	AGCTGGATGT	19530
	GCAGTGTCCA	GATGGTGACG	GTCCGGGGTG	AGGTCTCCAG	ACCCTGCGGT	GAGCTGGATG	TCCGCTGTCT	19600
55	GGATGGTGCA	GGTCTGGAGT	GAGGTCTGCCA	GGCCCTCGGT	GAGCTGGATG	TATGGAGTCC	GGATGGTGCC	19670
	GGTCCGGGGT	GAGGTCTGCCA	GACCTCTGCT	TGAGCTGGAT	GTGCGGTGTC	TGGATGGTAC	AGGTCTGGAG	19740
	TGAGGTGCGC	AGACCTGTCT	GTGAGCTGGA	TATGCGGTGT	CCGGATGGTG	CAGGTCAAGG	GTGAGGTCTC	19810
	CAGGCCCTCG	GTGAGCTGGA	GGTATGGAGT	CCGGATGGTG	CAGGTCTCCG	GTGAGGTCTC	CAGGCCCTCG	19880
	TGTGAACCTG	ATGTGCGGGC	TCTGGATGGT	GCAGGTCTGG	GGTGTGGTCC	CCAGGCCCTC	GGTGAGCTGG	19950
60	AGGTATGGAG	TCCGGATGAT	GCAGGTCCGG	GGTGAGGTGC	CCAGGCCCTG	CTGTGAGCTG	GATGTGCGGC	20020
	GTCTGGATGG	TGCAGGTCTG	GGGTGTGGTC	GCCAGGCCCT	CGGTGAGCTG	GAGGTATGGA	GTCCGGATGA	20090
	TGCAGGTCCG	GGGTGAGGTT	GCCAGGCCCT	GCTGTGAGCT	GGATGTGCTG	TATCCGGATG	GTGCACTCCG	20160
	GGGTGAGGTC	GCCAGGCCCT	GCTGTGAGCT	GGATGTGCTG	TATCCGGATG	GTGCAAGTCT	GGGTGAGGTC	20230
	CACCAGGCC	TGCGGTGAGC	TGGTTGTGCG	GTGTCCGGTT	GCTGCAAGTC	CGGGGTGAGT	TGCGAGGCC	20300
65	CTCGGTGAGC	TGGATGTGCG	GTGTCCCGGT	GTCCGGATGG	TGCAGGTCCA	GGGTGAGGTC	GCTAGGCCCT	20370
	TGGTGGGCTG	GATGTGCCGT	GTCCGGATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGGTGAGCTG	20440
	GATGTGCCGT	GTCTGCATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGGTGGGCTG	GATGTGTGGT	20510
	GTCCGGATGG	TGCAGGTCCG	GCGTGAGGTC	GCCAGGCCCT	GCTGTGAGCT	GGATGTGCGG	TGCTTGGATG	20580
	GTGCAAGTCC	GGGTGAGGTT	AGCCAAGGCC	TTCCGTGAGC	TGGATGTGGG	GTGTCCGGAT	GGTGCAAGTC	20650
70	CGGGGTGAGG	TGCGCAAGCC	CTGCGGTAG	CTGGATATGC	GGTGTCCGGA	TGGTGCAGGT	CCGGGGTGAG	20720
	GTCAACAGGC	CCTGCGGTTA	GCTGGATGTG	CGGTGTCTGG	ATGGTGACAG	TCCGGGGTGA	GGTCCGCCAG	20790
	CCCTGCTGTG	AGCTGGATGT	GCTGTATCCG	GATGTGTGAC	GTCGGGGGTG	AGGTCCGCCAG	GCCCTGCACT	20860
	GAGCTGGATG	TGCTGTATCC	GGATGGTGCA	GGTCTGGCGT	GAGGTCCGCA	GGCCCTGCGG	TTAGCTGGAT	20930
	ATGCGGTGTC	GGATGGTGCA	GGTCCGGGGT	GAGGTCCGCA	GGCCCTGCGG	TTAGCTGGAT	GTGCGGTGTC	21000
75	CGGATGGTGC	AGGTCTGGGG	TGAGGTGCGC	AGGCCCTGCT	GTGAGCTGGA	TGTGCTGTAT	CCGGATGGTG	21070
	CAGGTCCGGG	GTGAGGTCCG	CAGGCCCTGC	GGTGAGCTGG	ATGTGCTGTA	TCCGGATGGT	GCAAGTCTGG	21140
	CGTGAGGTCG	CCAGGCCCTG	CGGTGAGCTG	GATGTGCACT	GTACGGATGG	TGCAGGTCCG	GGGTGAGGTC	21210
	GCCAGGCCCT	GCGGTGGGCT	GTATGTGTGT	TGCTGAGATG	GTGCAAGTCC	GGGGTGAAGT	GCCAGGCCCT	21280

	TGCGGTGAGC	TGGATGTGTG	GTGCTGGAT	GCTGCAGGTC	CGGGGTGAGT	TGCGCAGGCC	CTCGGTGAGC	21350
	TGGATATGCG	GTGTCCCGCT	GTCCGAATGG	TGCAGGTCCA	GGGTGAGGTC	GCCAGGCCCT	TGGTGGGCTG	21420
	GATGTGCGGT	GTCCGGATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGGTGAAGCTG	GATGTGCGGT	21490
5	GTCCGGATGG	TGCAGGTCCG	GGGTGAGGTC	ACCAGGCCCT	CGGTGATCTG	GATGTGGCAT	GTCTTCTCG	21560
	TTTAAGGGGT	TGGCTGTGTT	CGGCCCGCAG	AGCACCGTCT	CGGTGAGGAG	ATCCTGGCCA	AGTTCTCTCA	21630
	CTGGCTGATG	AGTGTGTACG	TCGTCAAGCT	GCTCAGGTCT	TTCTTTTATG	TCACGGAGAC	CACGTTTCAA	21700
	AAGAACAGGC	TCTTTTCTA	CGGGAAGAGT	GTCTGGAGCA	AGTTGCAAAAG	CATTGGAATC	AGGTACTGTA	21770
	TCCCAACGCC	AGGCCTCTGC	TTCTCGAAGT	CCTGGAAACAC	CAGCCCGGCC	TCAGCATGCG	CCTGTCTCTCA	21840
10	CTTGCTCTGT	CTTCCCTGGC	TGTGCAGCTC	TGGGTGGGA	GCCAGGGGCC	CCGTACACAG	CCTGTCTCAA	21910
	GTGGATTCTG	TGCAAGGCTC	TGACTGCCTG	GAGCTCAGCT	TCTCTTACTT	GTAATAATCAG	GAGTTTGTGC	21980
	CAAGTGGTCT	CTAGGGTTTG	TAAAGCAGAA	GGGATTAAAA	TTAGATGGAA	ACACTACCAC	TAGCCTCTCT	22050
	GCCTTTCCCT	GGGATGTGGG	TCTGATTCTC	TCTCTCTTTT	TTTTTCTTTT	TTTGAGATGG	AGTCTCACCT	22120
	TGTTGCCAG	GCTGGAGTGC	AGTGGCATAA	TCTTGGCTCA	CTGCAACCTC	CACCTCTCTG	GTTTAAAGCGA	22190
15	TTACCCAGCC	TCAGCCTCCT	AAGTAGCTGG	GATTACAGGC	ACCTGCCACC	ACGCTCTGGT	AATTTTGTGA	22260
	CTTTTAGAGG	AGACGGGGTT	TCACCATGTT	GGCCAGGCTG	GTCTCGAACT	CATGACCTCA	GGTGATCCAC	22330
	CCACCTTGGC	CTCCCAAAGT	GCTGGGTTTA	CAGGCTAAGC	CACCGTGCCC	AGCCCCCGAT	TCTCTTTTAA	22400
	TTTCATGCTG	TCTGTATGAA	TCTTCAATCT	ATTGGATTAA	GGTCATGAGA	GGATAAAATC	CCACCCACTT	22470
	GGCGACTCAC	TGCAGGGAGC	ACCTGTGCAG	GGGATAGGCT	AGTTCCACCA	AGTTCCACCA	TGAGCTAATC	22540
20	TCTAGGTGGC	TGCATTGAA	TGGCTGTGAG	ATTTTGTCTG	CAATGTTCCG	CTGATGAGAG	TGTGAGATTG	22610
	TGACAGATTG	AAGCTGGATT	TGCATCAGTG	AGGGACGGGA	GCGCTGGTCT	GGGAGATGCC	AGCCTGGCTG	22680
	AGCCAGGCTC	ATGGTATTAG	CTTCTCCGTG	TCCCGCCAG	GCTGACTGTG	GAGGGCTTTA	GTCAAGAAAT	22750
	CAGGGCTTCC	CCAGCTCCCC	TGCACACTCG	AGTCCCTGGG	GGGCTTGTG	ACACCCCATG	CCCCAAATCA	22820
	GGATGTCTGC	AGAGGGAGCT	GGCAGCAGAC	CTCGTCAGAG	GTAACACAGC	CTCTGGGCTG	GGGACCCGGA	22890
25	CTGTGTGCTG	GGGCCAATTC	CTTGATCTGT	GGGGAGGGTC	AGGGCTTTCC	CTGTGGGAAC	AAGTTAATAC	22960
	ACAATGACCC	TTACTTAGAC	TTTACACGTA	TTTAATGGTG	TGCGACCCAA	CATGGTCATT	TGACCAAGAT	23030
	TTTGAAAGA	ATTTAATTGG	GGTGACCGGA	AGGAGCAGAC	AGACGTGGTG	GTCCCAAGA	TGCTCTTGT	23100
	CACTACTGGG	ACTGTTGTTT	TGCTTGGGGG	GCCTTGGAGG	CCCTCTCTCC	CTGGACAGGG	TACCTGTGCT	23170
	TTTCTACTCT	GCTGGGCTG	CGGCTCGCG	TCAGGGCACC	AGTCCGGAG	CACCCGGGCG	CCCACTGTCC	23240
30	ACGGAGTGCC	AGGCTGTGAG	CCACAGATGC	CCAGGTCCAG	GTGTGGCCGC	TCCAGCCCCC	GTGCCCCCAT	23310
	GGGTGGTCTT	GGGGGAAAG	GCCAAAGGCA	GAGGTGTGAG	GAGACTGGTG	GGCTCATGAG	AGCTGATTCT	23380
	GCTCCTTGGC	TGAGCTGCCC	TGAGCAGCCT	CTCCCGCCCT	CTCCATCTGA	AGGGATGTGG	CTCTTCTCAT	23450
	CTGGGGGCTC	TGCTTGGGGG	CAGCCTTGGG	CTACCCCAAG	GGCTGTACCA	GAGGGACAGG	CATCCTGTGT	23520
	GGAGGGGAT	GGGTTCACGT	GGCCCCAGAT	GCAGCCTGGG	ACCAGGCTCC	CTGTGTCTGA	TGGTGGGACA	23590
35	GTACCCCTGG	GGGTTGACCC	CGGGACTGGG	CGTCCCCAGG	GTTGACTATA	GGACCAAGTG	TCCAGTGCC	23660
	CTGCAAGTAG	AGGGGCTCTC	AGAGGCTCT	GGCTGGCATG	GGTGGACGTG	GCCCGGGCA	TGGCTTCTAG	23730
	CCTGTGCTGC	CGTGGGTGCC	CTGAGCCCTC	ACTGAGTCCG	TGGGGGCTTG	TGGCTTCCCG	TGAGCTTCCC	23800
	CGTAGTCTGT	TGCTTGGCTG	AGCAAGCCTC	CTGAGGGGCT	CTCTATTGCA	GACAGCACTT	GAAGAGGGTG	23870
	AGCTGCGGGG	AGCTGTGCGA	AGCAGAGGTC	AGGCAAGCAT	GGGAAGCCAG	GCCCGCCCTG	CTGAGCTCCA	23940
40	GACTCCGCTT	CATCCCCAAG	CCTGACGGGG	TGCGGCGCAT	TGTGAACATG	GACTACGTCT	TGGAGCCAG	24010
	AACGTTCCCG	AGAGAAAAGA	GGGTGGCTGT	GCTTTGGTTT	AATTCCTTTT	TTAAACAGAA	GTGCGTTTGA	24080
	GGCCCACTT	TGATATCAGC	TTAGATGAAG	GGCCCGGAGG	AGGGGCCACG	GGACACAGCC	AGGGCCATGG	24150
	CACGGCGCCA	ACCCATTGTG	GGCCACAGTG	AGGTGGCCGA	GGTGCCGGTG	CCTCCAGAAA	AGCAGCGTGG	24220
	GGGTGTAGGG	GGAGCTCCTG	GGGCAGGGAG	AGGCTCTGAG	GACCACAAGA	AGCAGCCGGG	CCAGGGGCTG	24290
45	GATGCAGCAC	GGCCCGAGGT	CCTGGATCCG	TGCTCTGCTG	TGGTGCCGAG	CCTCCGTGGC	CTTCCGCTTA	24360
	CGGGGGCCCG	GGACACAGGC	ACGACTGCCA	GGAGCCCAAC	GGGCTCTGAG	GATCCTGGAC	CTTGGCCCAT	24430
	GGCTCTCTGA	CCCCACCCCT	GTGGCTGCGG	TGGCTGCGGT	GACCCCGTCA	TCTGAGGAGA	GTGTGGGGTG	24500
	AGGTGGACAG	AGGTGTGGCA	TGAGGATCCC	GTGTGCAACA	CACATGCGGC	CAGGAACCCG	TTTCAACAG	24570
	GGTCTGAGGA	AGCTGGGAGG	GGTCTAGGTT	CCCCGGTCTG	GGTGGCTGGG	GACACTGGGG	AGGGGCTGCT	24640
50	TCTCCCTGG	GTCCCTATGG	TGGGTGGGCG	ACTTGGCCGG	ATCCACTTTC	CTGACTGTCT	CCCATGCTGT	24710
	CCCCGCCAGG	CCGAGCGTCT	CACCTCGAGG	GTGAAGGCAC	TGTTACGCGT	GCTCAACTAC	GAGCGGGCGC	24780
	GGCGCCCCCG	CCTCCTGGGC	GCCTCTGTGC	TGGGCTGGA	CGATATCCAC	AGGGCCTGGC	GCACCTTCTG	24850
	GCTGCGGTGT	CGGGCCCAAG	ACCCGCCGCC	TGAGCTGTAC	TTTGTCAAGG	TGGGTGCCGG	GGACCCCGCT	24920
55	GAGCAGCCCT	GCTGGACCTT	GGGAGTGGCT	GCCTGATTGG	CACCTCATGT	TGGGTGGAGG	AGGTACTCCT	24990
	GGGTGGGCCG	CAGGGAGTGC	AGGTGACCTT	GTCACTGTG	AGGACACACC	TGGCACCTAG	GGTGGAGGCC	25060
	TTTCAGCCTT	CCTGCAGCAC	ATGGGGCCGA	CTGTGCACCC	TGACTGCCCG	GGCTCCTATT	CCCAAGGAGG	25130
	GTCCCACTGG	ATTCCAGTTT	CCGTGAGAGA	AGGAACCCGA	ACGGCTCAGC	CACCAGGCCG	CGGTGCCCTT	25200
	CACCCCAAGT	CTGAGCCAGG	GGTCTCCTGT	CCTGAGGCTC	AGAGAGGGGA	CACAGCCCGC	CCTGCCCTTG	25270
60	GGGTCTGGAG	TGGTGGGGGT	CAGAGAGAGA	GTGGGGGACA	CCGCCAGGCC	AGGCCCTGAG	GGCAGAGGTG	25340
	ATGCTGTGAGT	TTCTGCGTGG	CCAATGTGAG	TCTCTCGCC	TCCACTCACA	CAGGTGGATG	TGACGGGGCG	25410
	GTACGACACC	ATCCCCAGG	ACAGGCTCAC	GGAGGTCATC	GCCAGCATCA	TCAAACCCCA	GAACACGTAC	25480
	TGCGTGCGTC	GGTATGCCGT	GGTCCAGAAG	GCCGCCCATG	GGCACGTCCG	CAAGGCCCTT	AAGAGCCACG	25550
	TAAGGTTAC	GTGTGATAGT	CGTGTCCAGG	ATGTGTGTCT	CTGGGATATG	AATGTGTCTA	GAATGCAGTC	25620
65	GTGTCTGTGA	TGCGTTTCTG	TGGTGGAGGT	ACTTCCATGA	TTTACACATC	TGTGATATGC	GTGTGGGCA	25690
	CGTGTGTGTC	GTGGTGCATG	TATCTGTGGC	GTGCATATTT	GTGGTGTGTG	TGTGTGTGGC	ACGTGTGTGT	25760
	CCATGGTGTG	TGTGCCCTGTG	GTGTGCATGT	GTGTGTGTCT	GTGACACGTG	CATGTTTATG	CTGTGTGCTG	25830
	CATGTCTGTG	ATGTGCCAT	TTGTGGTGTG	TGTGTGCATG	TGTCGGTGAC	ATATGCGTGT	CTATGGCATG	25900
	GGTGTGTGTG	GCCCTTGGC	CTTACTCCTT	CCTCCTCCAG	GCATGGTCCG	CACCATTTGT	CTCACGCTCT	25970
	CGGGTGCTGG	TTTGGGGAGC	TCCACATTCA	GGGTCTCTAC	TTCTAGCATG	GGTGCCCTG	TCTGTCTACA	26040
70	GGGCTGGGCC	TTGAGAGCTG	TAAGCCAGGT	TTGAGAGGAG	AGTAGGGATG	CTGGTGGTAC	CTTCTGGAC	26110
	CCCTGGCACC	CCCAGGACCC	CAGTCTGGCC	TATGCCGGCT	CCATGAGATA	TAGGAAGGCT	GATTACAGCC	26180
	TGCGTCCCCG	GGACACACTC	CTCCAGAGC	GGCCGGGGGC	CTTGGGGCTC	GGCAGGGGTT	AAAGGGGGCC	26250
	TGGGCTTGGG	TTCCACCCCA	GTGGTCATGA	GCACGCTGGA	GGGGTAAGCC	CTCAAAGTCG	TGCCAGGGCC	26320
	GGGTGACAG	GTGAAGAAGT	ATCCCTGGAG	CTTCGGTCTG	GGGAGAGGCA	CATGTGGAAA	CCCACAAGGA	26390
75	CCTCTTTCTC	TGACTTCTTG	AGCT					26414

Contig 2:

	TGTGGGATTG	GTTTTCATGT	GTGGGATAGG	TGGGGATCTG	TGGGATTGGT	TTTTATGAGT	GGGGTAACAC	70
	AGAGTTCAAG	GCGAGCTTTC	TTCTCTAGT	GGGTCTGCAG	GTGCTCCAAC	AGCTTTATTG	AGGAGACCAT	140
5	ATCTTCTCTT	GAACATATGGT	CGGGTTTATA	GTAAGTCAGG	GGTGTGGAGG	CCTCCCTTGG	GCTCCCTGTT	210
	CTGTTTCTTC	CACTCTGGGG	TCGTGTGGTG	CCTGCTGTGG	TGTGTGGCCG	GTGGGCAGGG	CTTCCAGGCC	280
	TCCTTGTGTT	CATTGGCCCTG	GATGTGGCCC	TGGCTACGCT	CCGTCTCTGG	AATTCCTCTG	CGAGTTGGAG	350
	GCTTTCCTTC	TTTCTTTTTT	TCITTCITTT	TTTTTTTTTT	TGATAACAGA	GTCTCGCTCT	TTTTTGCCCA	420
	GGCTGGAGTG	GTTTGGCGTG	ATCTTGGCTC	ACTGCAACCT	GTGCTTCCTG	AGTTCAAGCA	ATTCTCTTGC	490
10	CTCAGCCTCC	CAAGTAGCTG	GAATTATAGG	CGCCACCCAC	CATGCTGACT	AATTTTTGTA	ATTTTAGTAG	560
	AGACGAGGTT	TCTCCATGTT	GGCCAGGCTG	GTCTCGAACT	CCTGACCTCA	GGTGATCCTC	CCACCTCGGC	630
	CTCCCAAAGT	GCTGGGATGA	CAGGTGTGAA	CCGCCGCGCC	CGGCCGAGAC	TCGCTTCCTG	CAGCTTCCTG	700
	GAGATCTGCA	GCGATAGCTG	CCTGCAGCCT	TGGTGTGAC	AACCTCCGTT	TTCTTCTCTC	AGGTCTCGCT	770
	AGGGGCTCTT	CCATTTCATG	ACTCTCTTCA	CAGAAGAGTT	TCACGTGTGC	TGATTTCCCG	GCTGTTTCTT	840
15	GCGTAATTGG	TGTCTGCTGT	TTATCGATGG	CCTCCTTCCA	TTTCTTTTAG	GCTTTGTTTA	TGTTTGTCTT	910
	TCCGGCTCCT	TGAAGGAAAA	GTTTCGATTA	TGGATGTTTG	AACCTTCTTT	TCTAAACAAG	CATCTGAAGT	980
	TGCCGTTTTT	CCTCTAAAGC	AGGGATCCCG	AGGCCCTTGG	CTGTGGAGTG	GCACCGGTCT	GGGGCTGTTT	1050
	AGGAACCCCG	GCGACAGCGG	GAGGCTAGGT	GGGGTGTGGG	GAGCCAGCGT	TCCCGCTTGA	GCCCGCGCCC	1120
	TCTCAGATCA	GCAAGTGGCAT	GCGGTGCTCA	GAGGCGCACA	CACCTTACTG	AGAATCTGTG	GAGAGAGGGG	1190
20	TCTAGATTCT	GTGCTCCTTA	TGGGAATCTA	ATGCCCTGATG	ATCTGAGGTG	GAACCGTTTG	CTCCCAAAAC	1260
	CATCCCCCTC	CCCACTGCTG	TCCTGTGGAA	AAATCGTCTT	CCACGAAACC	AGTCCCTGGT	ACCACAATGG	1330
	TTGGGGACCC	TGTGCTAAAG	ACCTGCTTCA	GCAGCCTCTC	GTCACTGTTG	ATATATTGGC	TTTTCTGTGT	1400
	TGAGTCCAGA	ATAATTACGG	ATTTCCTGTA	TGCTTTCGCG	CGACCTCAGA	CCCATGGGCT	ATTTGTGGCG	1470
25	GTGTGCGCTG	CTCCTGGGTT	GGGAAGGGTG	CAGGCCCCAT	GTACCTTCTT	GTTACTGCCT	TCCAGGTTGG	1540
	TTCTCAGGGT	TGAATCGTAC	TCGATGTGGT	TTTAGCCAC	GGCCCTGCGC	CCAGCTCCTG	GGGGCTGGGG	1610
	AACATGCTGA	AGCACAGAGT	CACCGTGGCG	GTCTTTTGAT	GCCTCACAAG	CTCAGAGCCT	CCTGTGTCCG	1680
	TGTTAGTGTG	TGTACGTGTC	CTGCTCACAT	CCTGTCTTGG	GGACGCAAGG	GCTTAGCAGG	TCCCGTAGTA	1750
	AATGACAAGC	GTCTTGGGGG	AGTCTGCAGA	ATAGGAGGTG	GGGGTGGCGG	TCTCTCTCCC	CGCTCTTCAG	1820
30	ACTCTTCTCC	TGCCCTGTGCT	GTGGCTGCAC	CTGCATCCCT	GCAATCCCTC	CAGCACTGGG	CTGGAGAGGG	1890
	CCGGGAGCTC	GAGTGCCACT	TGTGCCACGT	GACTGTGGAT	GGCAGTCCGT	CACGGGGGTC	TGATGTGTGG	1960
	TGACTGTGGA	TGGCGGTTGG	TCACAGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTCTG	TGGGGTCTGA	2030
	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTCTG	TGGGGTCTGA	2100
35	TGTTGACTGT	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	2170
	TGTTGACTGT	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	2240
	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTCTG	TGGGGTCTGA	2310
	TGTGGTGGA	CTGTGGATGG	CGGTCTGTGG	GTCTGATGTG	GTGACTGTGG	ATGGCGGTCTG	TGGGGTCTGA	2380
	TGTGGTGGA	CTGTGGATGG	TGATCGGTCA	CAGGGGTCTG	ATGTGTGGTG	ACTGTGGATG	CGGGTCTGTG	2450
40	GGTCTGATGT	GTGGTGACTG	TGGATGGTGA	TGGGTCTGATG	GTGACTGTGG	ATGGCGGTCTG	TGGGGTCTGA	2520
	GTGCTGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTGG	GTCCCGGGGG	TCTGATGTGT	GGTGACTGTG	2590
	GATGGCGGATC	GGTCACAGGG	GTCTGATGTG	TGGTGACTGT	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2660
	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	TGGTGACTGT	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2730
45	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	TGGTGACTGT	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2800
	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	TGGTGACTGT	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2870
	GTGTTGGTGC	TGTGGATGGC	GGTCTGGGGT	TCTGATGTGT	GCTGACTGTG	GATGGCGGTCT	TGGGGTCTGTG	2940
	ATGTGTGGTG	ACTGTGGATG	GCGGTCTGTG	GGTCTGATGT	GGTACTGTG	GATGGCGGTCT	GTGGGTCTGTG	3010
	ATGTGTGGTG	ACTGTGGATG	GCGGTCTGTG	GGTCTGATGT	GGTACTGTG	GATGGCGGTCT	GTGGGTCTGTG	3080
50	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3150
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3220
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3290
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3360
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3430
55	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3500
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3570
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3640
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3710
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3780
60	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3850
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3920
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	3990
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4060
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4130
65	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4200
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4270
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4340
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4410
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4480
70	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4550
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4620
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4690
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4760
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4830
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4900
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	4970
	GGGTCTGATG	TGTGGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	5040

	CTGCCCGGGG	CCACCAGAGT	CTCCTTTTCT	GGCCCCCGCC	CCCTCCGGCT	CCTGGGCTGC	AGGCTCCCCA	5110
	GGCCCCGGAA	ACATGGCTCG	GCTTGCGGCA	GCCCGAGCGG	AGCAGGTGCC	ACACGAGGCC	TGGAATGGC	5180
	AAGCGGGGTG	TGGAGTTGCT	CCTGCGTGGA	GGACGAGGGG	CGGGGGGTGT	GTCTGGGTCA	GGTGTGCGCC	5250
5	GAGCGTTTGA	GCCTGCAGCT	TGTGAGCTCC	AAGTTACTAC	TGACGCTGGA	CACCCGGCTC	TCACACGCTT	5320
	GTATCTCTCT	CTCCCCGATC	AAAAAGGATT	TATCCGATTC	TCATTCTCTG	CCCTGTCTGT	TGACCCCGCG	5390
	GAGGGCGCGG	GCTCTTCTCT	CTGTGACTAG	ATTTCCCATC	TGGAAAGTGC	GGGGTTGACC	GTGTAGTTTG	5460
	CTCCTCTCGG	GGGGCCTGTG	GTGGCCATGG	GGCAGGCGGC	CTGGGAGAGC	TGCCGTACAC	CAGCCACTGG	5530
	GTGAGCCACA	CTCACGGTGG	TAGAGCCACA	GTGCCCTGGT	CCACATCAGG	TCCTCTGGAT	TTTAAAGTAA	5600
10	ACCACACACC	TCCCGGCGAG	CATCTGCTCG	CGACCTCTGT	TGTGCCCTGG	GAGAGTGGTA	GCACGGAGGA	5670
	AATTCGTGCA	CACCTCAAGG	CATCAGCAAG	GTACCTCCGA	GTCAGGTGGA	ACGTGGAGGC	CTCTCTCTGG	5740
	GATCGTCTCC	AGCGGATAAA	GGACTGTGCA	CAGCTTCGGA	AGCTTTTATT	TAAAAATATA	ACTATTAAAT	5810
	ATTGCAATTAT	AAGTAATCAC	TAATGGTATC	AGCAATTATA	ATATTATTAT	AAGTATAATT	AGAAATATTA	5880
	AGTAGTACAC	ACGTTCTGGA	AAAACACAAA	TTGCACATGG	CAGCAGAGTG	AATTTTGGCC	GAGGGACACG	5950
15	TGTGCACATG	TGTGTAAGCG	CCCCCAGGCG	CCACAGAATT	CGCTGACAAA	GTCACCTCCC	CAGAGAAGCC	6020
	ACCACGGGGC	TCCTTCGTGG	TCGTGAATTT	TATTAAGATG	GATCAAGTCA	CGTACCGTCC	ACGTGTGGCA	6090
	GGGCTTTGGG	GAATGTGAGG	TGATGACTCG	GTCTCATGCG	CCTGACAGAC	AGGAGGTGAC	TGTGTCTGTC	6160
	CTGTCCCTAG	GACACGGACA	GGCCCGAAGC	TCTAGTCCCC	ATCGTGGTCC	AGTTTGGGCT	CTGAATAAAA	6230
	ACGTCTTCAA	AACCTGTTGC	CCCAAAACT	AAGAACAGAG	AGAGTTTCCC	ATCCCATGTG	CTCACAGGGG	6300
20	CGTATCTGCT	TGCGTTGACT	CGCTGGGCTG	GGCCGACTCC	TAGAGTTGGT	CGGTGTGCTT	CTGTGCAAAA	6370
	AGTGCACTCC	TCTTGCCCAT	CACTGTGATA	TCTGCACCAG	CAAGGAAAGC	CTCTTTTCTT	TTCTTTCTTT	6440
	TTTTTTTTTT	GAGACGGAAC	GTCACCTGTT	TCTGCTGGG	CTTGAGTGCA	GTGGCGCGAT	CTCAACTCAC	6510
	TGCAACCTCC	GCCTCCCGGG	TTCCAGCATT	TCTCTGCCT	CAGCCTCCCC	AGCAGCTGAG	ATTACAGGCA	6580
	CCACCCCTCT	GGCCTCTGGT	AATTTTGTGA	TTTTTAGTAG	AGAGGGGTTC	TTGCCATGTT	TGCCAGGCTG	6650
25	GTCTCGAAGT	CCTGACCTCA	GGTGATCCAC	CCACCTCGGC	CTCCCAAAGT	GCTGGGATTA	CAGGTGTGAG	6720
	CCATCACGCC	CAGCCGGAAA	GCCTCTTTTT	AAGGTGACCA	CCTATAGCGC	TCCCGAAAA	TAACAGGCTC	6790
	TGTTTTTGCA	GTAGGCTGCA	AGCGTCTCTT	AGCAACAGGA	GTGGCTCTCT	GTGGGCTCTG	GGGATGGCTG	6860
	AGGGTCTCGT	GGCAGCCATG	CCTTCTGTGT	GCACCTTTAG	GTTCACCGGG	GCTATTCTGC	TCTCACTGTT	6930
	TGTCTGAAAA	CGCACCCCTG	GCATCCCTTG	TGGAGAGTTC	TCTGCTCTCT	GTTGGTCAAT	CTGAAACTAG	7000
30	GGGCAAGGTT	GTATCCGTTG	GGCGCGAGCG	GTCACATGTA	GGGTCAATGAG	TCTTTCACCG	TGGACAAATT	7070
	CCTTGAAAAA	AAAAAAAGGA	TTCGGGTTAA	GCATTTCATC	CGGGTCAAGT	GTCTGGTTCT	GTGAATAAAC	7140
	CTTAAGATT	AAGAAACCTT	AATGAAAGAA	AACCTTGATG	ATTGAGAGCA	AGGATGTGGT	CACACCTGTG	7210
	GCTGGATCTG	TTTCAAGCCG	CCCACTGCTT	GGTGAGAGTG	GGGAGCAGGG	ATTGTTTGT	CAGAGGCTCT	7280
	ATCTGGTATG	TTTCTGAGGT	GTTCGCCGCG	TGAATGGTAG	ACGTGTCGTT	TGTGTGTATG	AGGTTCTGTG	7350
35	TCTGTGTGTG	GCTCGGTTTG	AGTGTACGCA	TGTCCAGCAC	ATGCCCTGCC	CGCTCTCTAC	CTGTGTCTTC	7420
	CCGCCCCAGG	TCCTACGTTC	AGTCCGAGGG	GATCCCGCAG	GGCTCCATCC	TCTCCACGCT	GCTCTGCGAC	7490
	CTGTGCTACG	CGGACATGGA	GAACAAGCTG	TTTGGCGGGA	TTCCGGCGGA	CGGGTGAGGC	CTCCTCTTCC	7560
	CCAGGGGGGG	TTGGGTGGGG	GTGATTGTC	TTTTGATGCA	TTGAGTGTTA	ATATTCTGTG	TGCTCTGGAG	7630
	ACCATGACTG	CTCTGTCTTG	AGGAACAGCA	CAAGGTTGCA	GCCCTTCTTT	GGTATGAAG	CGCAGGGGAG	7700
40	GGGTTGACAC	GCCTGAGGCT	TGCGGGCTCC	ACCGAGGCTC	TGTCAGCGG	CCATGTCCAG	AGGCTCTAGG	7770
	GCTCAGCAGG	CGGAGGCGCC	GCTGCCCTGC	ATGATGAGCA	TGTGAATTCA	ACACCGAGGA	AGCACACCAG	7840
	CTTCTGTGAC	GTCACCCAGG	TTCCGTTAGG	GTCCCTGGGG	AGATGGGGGT	GGTGAGCCCT	GAGGCCCCAC	7910
	ATCTCCACGC	AGGCCCTCGA	CAGGTGGGCT	GGACTGGGCG	CCTCTTCAGC	CCATTGCCCC	TCCCACTTGC	7980
	ATGGGGTCTA	CACCCAAAGGA	CGCACACACC	TAAATATCGT	GCCAACTTAA	TGTGGTTCAA	CTCAGCTGGC	8050
45	TTTTATTGAC	AGCACTTACT	TTTTTTTTTT	TAATACTTTA	AGTTCTAGGG	TACATGTGCA	CGACGTGACG	8120
	GTATAGTTACA	TATGTATACA	TGTGCCATGT	TGGTGTGCTG	CACCCATTAA	CTCATCATTT	ACATTAGGTA	8190
	TATCTCCTAA	TGCTATCCCT	CCCCACTCCC	CCCATCCCAT	GACAGGCCCT	GGTGTGTGAT	GTTCCTCCAC	8260
	CTGTGTCCAA	GTGTTCTCAT	TGTTCACTGT	CCACCTGTGA	GTGAGAACAT	GTGGTGTGTT	GTTCCTTTTC	8330
	CTTGCAATAG	TTTGCTCAGA	GTGATGGTTT	CCAGCTTCGT	CCATGTCCCT	ACAAAGGACA	TGAATCTATC	8400
50	CTTTTCTCAT	TAAGGTTCAA	ATTCCGTGGT	GTATATGTGC	CACATTTTCT	TAATCCAGTC	TATCATCGAT	8470
	GGACATTTGG	GTGGTTGCA	AGTCTTTGCT	ACTGTGAATA	GTGCCGCAAT	AAACATACGT	GTGCATGTGT	8540
	CTTTATAGCA	GCATGATTTA	TAATCTTTTG	GGTATATACC	CAGTAATGGG	ATGGCTGGGT	CAATGGGTAT	8610
	TTCTAGTTCT	AGATCCTTGA	GAATCACCA	CACGTCTTTC	CACAATGGTT	GAACATGTTT	ACACTCCAC	8680
	CAACAGTGTA	AAAGTGTCTT	GGTGCTGGAG	AGGATGTGGA	CAGCAGTTAT	TTTTTATGTA	AAATAGTATC	8750
55	ACTGAACAAG	CAGACAGTTA	GTGAAGGATG	CGTCAGGAAG	CCTGACGGCC	ACACAGCCAT	TTCTCTCGAA	8820
	GACTCCGGGT	TTTTCTGTG	CATCTTTTGA	AACCTAGACT	CCAATTATAG	CATGTACAGT	GGATCAAGGT	8890
	TCTTCTTCAT	TAAGGTTCAA	GTTCAGATT	GAAATAAGTT	TATGTAACAG	AAACAAAAAT	TTCTGTGACA	8960
	CACAACCTGC	TCTGGGATTT	GGAGGAAAGT	GTCCCTGAGC	TGGCGGCACA	CTGGTCAGCC	CTCTGGGACA	9030
60	GGATACCTCT	GGCCCATGGT	CATGGGGCGC	TGGGCTTGGG	CCTGAGGGTC	ACACAGTGCA	CATGCCCCAG	9100
	CTTCTCTGTG	ATAGGATCTG	GGTCTCGGAT	CATGCTGAGG	ACCACAGCTG	CCATGCTGGT	AAAGGGCACC	9170
	ACGTGGCTCA	GAGGGGGCGA	GGTTCCGAGC	CCGAGCTTTC	TTACCTGCTT	CAGTTATTTT	TCCCTAAGAG	9240
	TCTGAGAAAT	GGGGCCGCGC	CTGATGGCCT	TCGTTCTGCT	TCAGCTGGCA	CAGAAATTGCA	CAAGCTGATG	9310
	GTAACACTGT	AGTACTTATA	ATGAATGAGG	AATTGCTGTA	GCAGTTAACT	GTAGAGAGCT	CGTCTGTTGG	9380
	AAAGAAATTT	AAGTTTTTCA	TTTAAACCGT	TTGGAGAAATG	TTACTTTTAT	TATGGCTGTG	TAAATTGTTT	9450
65	GACATTACAT	CCCTCGTAGA	CAGATACTAC	GTAAGAAAGT	TAAAGTTAAC	CTTGCTGTGT	ATTTTCCCTT	9520
	ATTTTAGGCT	GCTCTCGCTG	TTGTTGGATG	ATTCTTGTGT	GGTGACACCT	CACCTCACCC	ACCGGAAAAAC	9590
	CTTCTCTCAG	TGAGGCCCGT	GGCGTGTGTC	TGTGGGGACC	TCCACAGCCT	GTGGGCTTTG	CAGTTGAGCC	9660
	CCCCGTGTCC	TGCCCTTGGC	ACCGCAGCGT	TGTCTCTGCC	AAGTCTCTCT	TCTCTGCGCG	TGCTGGATCC	9730
	GCAAGAGCAG	AGGCGCTTGG	CCGTGCACCC	AGGCTTGGGG	GCGCAGGGGC	ACCTTCGGGA	GGGAGTGGGT	9800
70	ACCGTGCAGG	CCCTGGTCTT	GCAGAGACGC	ACCCAGGTTA	CACACGTGGT	GAGTGCAGGC	GGTGACCTGG	9870
	CTCCTGCTGC	TCTTTGGAAA	GTCAAGAGTG	GCGGCTCCTG	GGGCCCCAGT	GAGACCCCCA	GGAGCTGTGC	9940
	ACAGGGCCTG	CAGGGCCGAG	GGCGCAGCCT	CCTCCCCAGG	GTGCACCTGA	GGCTGCGGAG	AGCAGGAGCT	10010
	CTGAGGTGAG	CTGGCCACAC	CGGTCACGTT	CCTGCTGGGG	GTTGTTTGGG	ATCGGTGGGA	ATCGGTGGGA	10080
	GAATTTGGAT	TTGCTGAGTG	CTGCTGTCTT	GAACACCGGA	GATGGCTAGG	AGTGGGTTTC	AGAGTTGATT	10150
	TTTGTGAATC	AAACTAAAAT	CAGGCGACAG	GGACCTGGCC	TCAGCAGAGG	GGATTGTCCA	ATGTGGTCCC	10220
75	CCTCAAGGGC	GCCCCACAGA	GCCGCTGGGC	TGTGTTTAAA	GTGCGATTGT	ACGAGGGGAC	AGAAACCTTG	10290
	AAAGCTGTAA	AGGGAACCTT	CAGAAAATGT	GGCCGCCAGG	GGTGGTTTCA	GGTGCTTTGC	TGGGCTGTGT	10360
	TTGTGAAAAA	CCATTTGGAC	CGCCCTTCCA	AGTCCACCTT	CCAGGTCCAC	CCTCCAGGGC	CGCCCTGGGC	10430

	TGGGGGTATG	CCTGGCGTTC	CTTGTGCCGC	AGCCCCGAGC	ACAGCAGGCT	GTGCACATTT	AAATCCACTA	10500
	AGATTCACTC	GGGGGGAGCC	CAGGTCCCAA	GCAACTGAGG	GCTCAGGAGT	CCTGAGGCTG	CTGAGGGGAC	10570
	AGAGCAGACG	GGGAACGCTG	CTTCTGTGTG	GCAAGTTTCT	GAGGGTGCTG	GCCAGGGAGG	TGGCTCAGAG	10640
	TGTATGTGGG	GGTCCCACCG	GGGGCAGAAC	TCTGTCTCTG	ATGAGTCGGC	AGCCATGTAA	CAGGAAGGGG	10710
5	TGGCCACAGG	GAGCTGGGAA	TGCACCAGGG	GAGCTGCCGA	GCTGGCCGAG	GTCCACAGGC	CAGGCCACAG	10780
	GAAGGGCAGG	GGGACGCCCG	GGGCCACAGC	AGAGGCCGCA	GGAGGGGAAG	GGGATGCCCA	GGCCAGAGCA	10850
	GAGGCTACCG	GGCACAGGGG	GGCTCCCTGA	GCTGGGTGAG	CGAGGCTCAT	GACTCGGCGA	GGGAACCTCC	10920
	TTGACGTGAA	GCTGACGACT	GGTGTGGCCC	AGCTCACAGC	CCAGCCAGGT	CCCGCGCCTG	AGCAGGAACCT	10990
	CAGAACCCTC	CCCTTTGTCT	AAAGCACAGC	AGATGCCTTC	AGGGCATCTA	GGAGAAAACA	GGCAAAAGTCG	11060
10	TTGAGAAACG	TCTTAAAGA	AGGTGGGATG	GTGGCAATTT	CTTGTCCAGA	TTTTAGTCTG	CCCCGACCA	11130
	CAGATGAGTC	TATAACGGGA	TTGTGGTGT	GCCATGGGGA	CACATGAGAT	GGACCATCAC	AGAGGCCACT	11200
	GGGCTGCAC	CTCCCATCTG	AGTCTGGCT	GTCCCGGGTC	CAGGCCAGGT	TCTTGCATGC	TCACCTACCT	11270
	GTCTTGGCCG	GGAGACAGGG	AAAGCACCCC	GAAGTCTGGA	GCAGGGCTGG	GTCCAGGCTC	CTCAGAGCTC	11340
	CTGCCAGGCC	CAGCACCCCTG	CTCCAAATCA	CCACTTCTCT	GGGGTTTTCC	AAAGCATTTA	ACAAAGGGTGT	11410
15	CAGGTTACCT	CCTGGGTGAC	GGCCCCGCAT	CCTGGGGCTG	ACATTGCCCC	TCTGCCTTAG	GACCTGTGTC	11480
	CGAGGTGTCC	CTGAGTATGG	CTGCGTGGTG	AACTTGGCGA	AGACAGTGGT	GAACTTCCCT	GTAGAAGACG	11550
	AGGCCCTGGG	TGGCACGGCT	TTTGTTCAGA	TGCCGGCCCA	CGCCCTATT	CCCTGGTGGC	GGCTGCTGCT	11620
	GGATACCCGG	ACCCCTGGAGG	TGCAGAGCGA	CTACTCCAGG	TGAGCCGACC	TGGCCGGGAG	TGGAGCCTGT	11690
	GGCCGGCTGG	GGCAGGTGCT	GCTGCAGGGC	CGTTCGCTCC	ACCTCTGCTT	CCGTGTGGGG	CAGGCCGACTG	11760
20	CCAATCCCAA	AGGCTCAGAG	GCCACAGGGT	GCCTCTGCTG	CCATCTGGGG	CTGAGCAGAA	ATGCATCTTT	11830
	CTGTGGGAGT	GAGGGTGCTC	ACAACGGGAG	CAGTTTTCTG	TGCTATTTTG	GTAAAAGGAA	ATGGTGCACC	11900
	AGACTGGGT	GCAGTGAAGT	GTCTTCAGAA	AGCAGTCTGG	ATCCGAACCC	AGACGCCCGC	GGCCCTGCTG	11970
	GGCGTGAGTC	TCTCAAAACC	GAACACAGGG	GCCTTGCTGG	GCATGAGTCC	CTCTGAACCC	GAGACCTGCG	12040
25	GGCCCTGCTG	GGCGTGAGTC	TCTCCGAACC	CAGAGACTTC	AGGGCCCTTT	TGGGCGTGAG	TCTCTCCGCT	12110
	GTGAGCCCCA	CACCTCCAAGG	CTCATCCACA	GTCTACAGGA	TGCCATGAGT	TCATGATCAC	GTGTGACCCA	12180
	TGAGGGGACA	GGGCCATGGT	GTGGGGGGGG	TCTCTACAAA	ATTTCTGGGG	CTTGTTCCTC	CAGAGCCCCA	12250
	GAGCTCAAGG	CCCCGTCTCA	GGCTCAGACA	CAAAATGAAT	GAAGATGGAC	ACAGATGCAG	AAATCTGTGC	12320
	TGTTTCTTTT	ATGAATAAAA	AGTATCAACA	TTCCAGGCAG	GCGAAGGTGG	CTCACACCTA	TAATCCCAGC	12390
30	ACTTTGGGAG	GCCGAGGTGG	GTGGATCACT	TGAGGCCAGG	AGTTTGAGGC	CAACCTAACC	AAATAGTGA	12460
	AATTCCATTT	CTACTTAAAA	AATACAAAAA	TTAGCCTGGC	CTGGTGGCAC	ACGCCCTGTAG	TCCCGCTAT	12530
	CGCGGAGGCT	GAGGCAGGAG	AATCATTGGA	ACCCAGGAGG	CAGAGGTTGC	AGTGAGCCGA	GATCACACCA	12600
	CTGCACCTCA	GCGCTGGGCA	CAGAGTGAGA	TTTCATCTTA	AAAAAAGTATC	AAAAAGTATC	GATGTACCCA	12670
	AACCATAGTG	GACAGGTGTT	TTTTTATTCT	GTCCCTCGAT	AATATTACT	GGTGCTGTGC	TAGAGGCCCG	12740
35	AACTGGGGGT	GGCTTCTCTC	GAAGGGCACA	CCTTCATGGG	AAGAGAAATA	AGTGGTGAAT	GGTTGTTAAA	12810
	CCAGAGGTTT	AAACTGGGGT	CCTGTCTGTC	TGAGTTAACA	GTCCAGATCT	GGACTTTGCC	TCTTCCAGA	12880
	ATGCTCCCTG	GGGTTTGCTT	CATGGGGGAG	CAGCAGGTGT	GSACACCTCC	GTGATGGGGG	AGCAGCAGGT	12950
	GCAGACGCCC	TCATGATGGG	GGAGTGGCAG	GTGCAGACAC	CCCTGTGCTC	GGTGCCGAGC	ATGTCCTGCT	13020
	TGCAGCTCCC	TCCCCACAAG	GATGCCGGTC	TGCTGTGCTC	CCCAAGTCC	CTGTTTCCCT	CTCACAGGCT	13090
40	TACCTGGTCC	TGGCCTCCAC	TGGCTTTGTC	TGCATGATTT	CCACATTTCC	TGGGCTCCCA	GCACCTCTTC	13160
	GGCTCTCCCA	GGCACCTCTG	CAGTGTGGG	CATACAGTGC	AGCTGTGAAC	TGTCCACTGC	TTATTTTGTG	13230
	CCCCATGAAA	TGTATTTTTT	AGGACAGGCA	CCCTGTGTTT	CAGCCTCTGG	CACAGCATCA	GTGAATGTTA	13300
	TTGAAGGACA	AAGGACAGAC	AAACAATCA	GGAAAATGGG	TTCTCTCTAA	ACACATTGCA	AAGCCACAGA	13370
	GGCTAGTGCA	GGATGGGTGG	GCATCAGGTC	ATCAGATGTG	GGTCCAAATG	CAGAAATATC	TGTGCTCCCA	13440
45	AAGGCCACTT	GGTCAGAGTG	TGTGCTTGCA	GAGGTGGCTC	TAAAAGCTCA	GCAGTGGAGG	CAGTGGTTCG	13510
	CCATACTCAG	GGTGAATCTA	CATCTCTGT	GTCTGAAATA	TACAGCAGAG	GCTTGAAGGG	CTCTGGGAG	13580
	AAGAAAACAG	GCAAAATGAT	TAAGAAAAGT	GAAAAGGAA	AAGTGGTAAG	ATGGGAATTT	TCTTGTCCAG	13650
	ATTTTAGTCT	CCCAAAACCAC	AGCTCAGATG	GTAGAATGTG	GTCAGAACTG	ATGGACAGAA	CAATAGAAC	13720
	AAACGGGAAGC	CCTATCTCTC	AGAAAAGTGT	GTAAATGTGG	TATGTGGCAC	AGCTGATGGA	AAAGAGAGTG	13790
50	TGTGTGTAAT	TTTTTTTTCT	GAGAAAAGTG	ACTGGAAGCA	AATAAGTTGT	GTCTTTACAG	CATATACCAG	13860
	AGCAGATTCT	AGGTAGAAGA	GGAGACACAT	GCAAAACACA	CCAGCAACAG	AAATAAAACA	AAAGACTCAA	13930
	AGGGGAAGGA	GGTGAACGTT	CCCTGGTTTG	GTGTTGGGGA	AGGACACACA	GGGAGGGCGA	TGAAACCACT	14000
	GAGGCAACGG	GCAATTGCTT	CAGTCGACAG	AAACTCAGCT	TGCCTGAGCC	ACAGTGAAAA	TGGCCATTCC	14070
	CTGGAGCGTT	TGTGCACGTG	ATTTATTTAA	GGCGCCCTGT	GAGGTCTGTC	ACATTCTATC	TCTCACTTTG	14140
55	TTCTCTTAAC	CACCTGAGAG	GTAGAGGAGG	AAAGGCTCCA	GGGAGGACGC	CGCCCTTGGT	CACCCAGCTG	14210
	GCAAAAGGCA	TGCATGATTG	CAGCCTGGCC	TCTGTCTCCG	GGGCCCTTGC	TCTGCCCGAG	GACCCACAC	14280
	AAGTCAGACC	CATAGGCTCA	GGGTGAGCCG	GAGCCCAAGC	TCGTGTTGGG	GATGGCTGTG	AAAGAAAGAA	14350
	TGGACGTCTG	ATGCACACTT	GGGAAGGTCC	TACCAGCAGC	GTCAAAGAAA	TGCATGTGAA	ACTGACAGCG	14420
	AGACCCATCC	CTCAAAGAAA	CGCACGTGAA	ACTGATGGCG	AGACCTGTCC	CCATCCCTCA	TGCTGGCTCC	14490
60	TTTTCTGGGC	TTGCCAAGAG	CCAGCATCAG	GTTGAGGCAA	GCTGGAAAGA	CTTTTCTGGA	AAGCAGCTTG	14560
	TTTGCATGGA	AGTCCTCACA	ATGTCTCTGT	TCTTCCAGT	AATTCCACTT	CTGAAGTGAC	CAGACATTAT	14630
	CAGGGTCTTT	ATTTACCAT	TCCAGTGTTC	CAGGCAGGGG	GACTTGCCAC	AGCAAGTCAC	GAACTGTCCC	14700
	AAATACAGGG	CTAAGGAGAT	ATTATGCATC	ACAAAACCTT	CTCTGCCATT	AAACATTTTT	CAAGAATTTT	14770
	TTGAAGAATG	TTTAATGGCA	CAAAACGTTT	ATTTCAATGT	AGCAGTGTTC	AAAGCTGGAT	GTAAAAGAAC	14840
65	ACACCCAGG	AGCCTGCCGT	GAATGTCATG	TGTGTTTATC	TTTGAGCATG	GACATACATG	GGCAGTGAGT	14910
	GGTGGTGAGG	CCCTGGAGGA	CATCGGTGGG	ATGCCCTCCAT	CCTGCCCTCC	TGGAGACACC	ATGTGTGCCA	14980
	CGTGCACTCA	CTGGAGCCCT	GTTTAGCTGG	TGCCACCTGG	CTCTTCCATC	CCTGAGATTC	AAACACAGTG	15050
	AGATTCCCCA	CGCCCAACTC	AGTGTCTCTC	CACAAAAAAC	CTGAGTCACA	CCTGTGTTCA	CTCGAGGGAC	15120
	GCCCGGGAGC	CAGGGCTCCA	CAGTTTATTA	TGTGTTTTTG	GCTGAGTTAT	GTGCAGATCT	CATCAGGGCA	15190
70	GATGATGAGT	GCACAAACAC	GGCCGTGCGA	GGTTTGATA	CACCTCAACAT	CACCTAGCCAG	GTCCGTGGTG	15260
	AGTTTGGTCA	TGCAGAGTCT	GGATGGCATG	TAGCATTTGG	AGTCCATGGA	GTGAGCACCC	AGCCCCCTCG	15330
	GGCTGCAGCG	CATGCCCCAG	GCAGGACAAG	GAAGCGGGAG	GAAGGCAGGA	GGCTCTTTGG	AGCAAGCTTT	15400
	GCAGGAGGGG	GCTGGGTGTC	GGGCAGGCAC	CTGTGCTGA	CATTCCCCC	TGTGTCTCAG	CTATGCCCGG	15470
	ACCTCCATCA	GAGCCAGTCT	CACCTTCAAC	CGCGGCTTCA	AGGCTGGGAG	GAACATGCGT	CGCAAACTCT	15540
75	TTGGGGTCTT	GCGGCTGAAG	TGTTCACAGC	TGTTCTTGGA	TTTGACGGTG	AGCAGGCTGA	TGGTCAGCAC	15610
	AGAGTTTCAGA	GTTCAGGAGG	TGTGTGCCCA	AGTATGTGTG	TGTGTGTGTG	CGCGCGTGCC	TGCAAGGCTG	15680
	ATGGTGACTG	GCTGCACGTA	AGATGTCACA	TGTACGCATA	TACACGTGAG	CACATACATG	TGTGCATGTG	15750
	TGTACATGAA	GGCATGGCAG	TGTGTGCACA	GGTGTGCAAG	GGCACAAAGT	TGTGCACATG	CGAATGCACA	15820

	CCTGACATGC	ATGTGTGTTT	GTCACAGTC	GTGTGGGCAT	TCACGTGAGG	TGCATGCCGT	TGGGTGTGCA	15890
	CTGTGAGTAG	CTCTGTGTGA	CATAACATGT	ATTGAGGGGT	CCTCGTGTC	ATCCCCTGAT	GTCTCCAGCA	15960
	CCAGTGGCAC	TACTTACAGG	ATGAGACGGT	GTGCCAGGCC	TGTGTGGGCT	GAGGCTCTGA	AGCTCGAGCC	16030
5	CTCAGGGCAT	TGTCCCCATCT	GGGCATCCGC	GTCACATCCC	CTCTCGTGGG	GCTTCTGTGT	CCACTCCCCC	16100
	TCTCTGTGGG	CGATTACAT	CGACTCCACT	CCCTCTCTCC	TGTGGGCAT	CCGGTCCAAT	CCCCCTCTCT	16170
	GTGGGCATCT	CGGTCCAACT	CCCTCTCTGT	TGGGCATTGT	CGTCCATCCC	CTCTCCCTGT	TCTTCTCTGT	16240
	CTTGGCCGAG	CCTCGGGGGC	AGGCAGATGA	CACAGAGTCT	TGACTCTGCC	AGGGTGGTTC	CGACGTGCCG	16310
	GGTGAGGGCC	AGGCCGGATT	TCACTGGGAA	GAGGGATAGT	TCTTGTCAA	AATGTTCTCT	TTTCTTGTTC	16380
	CATCTGAATG	GATGATAAAG	CAAAAAGTAA	AAACTTAAAA	TCCCAAGAGT	GTTTCTACCC	TTTTCTCACT	16450
10	TTTCTTGGCG	ACTCTAGATG	AACAGCCCTC	ACGAGGTTGT	CACCACATAG	TACAAGATAG	TTCTCTGCA	16520
	GGCGTACAGG	TGAGGCCGCA	CCAAAGGGTG	CAGGCCCGAC	CTCCAGGAGC	CTCTCCCGCT	GTGCTCACTT	16590
	CTACCCCGGG	GCTTCACTTT	GGAACTCTTG	GGTTTATAGG	GCAAGGAATG	TCTTACCTGT	TCAGTGTGTG	16660
	TGCTGCCCTGT	GCACAGTTCT	GTTCCGCTGG	CTCTGTGCAA	AGCACCTGTT	CTCCATCTCT	GGGTAGTGGT	16730
	AGGAGCCGGT	TGGGCCCCAG	GTGTCCCCAC	TGTGCGTGTG	CACCTGGGCT	GGGACCTCAT	GGAGGCCCAT	16800
	CACGGGCGAG	AGGGGACATG	AGGGAAGAAG	TGTTTATGGG	GAGTCTTTAG	AGGAGGGGCT	GGGAAGGTGT	16870
	CTGAACAGTA	GATGGGAGAT	CAGATAGCCG	GAGGATTTGG	GGTCTCAGCA	AGGAGGCGCG	AGGTGGGTGT	16940
	AGGTGAGGGT	CGCTGGCCCC	ACCCCGCCAG	AGGTTGACGA	GAGCTGTGAG	TCCCCAACGC	CGCCGGCCAG	17010
	CACCTGTGCT	CTGGGCATGC	CTGTGCTCT	GGAACTPTCC	CTGTCTGGC	TGGTCAGGGG	GTGCCCTCTG	17080
20	CAGAATTCGA	CAACTTTATC	ACAGAGGGAA	GGGCCAATCT	GTGGGAGCCA	CAGGGCCAGC	TTTGCCTCTG	17150
	AGTCAGGGCA	GGTGGTGGCA	GGAGCTCCGG	GGCTGTACCA	AGGGGACATG	GCGCCACCAC	GGCCCGGGCC	17220
	TCCACCTCAA	CAGGCCCTCCC	GAGCCTAGTT	GAGCTGAATG	CAGGAGGGCC	GAAGCCCTCG	CCCATAGAGT	17290
	GCTGAGAAAG	AGTGTGAGCA	TTTGTGTTAC	CCAGGGCCGA	GCTGAGCCCA	ATTACCGTGC	ACACTTGAAT	17360
	TGAATATGAG	TGCTGCTCTA	TCGTGTGAAC	CCAGCAAGGG	CTCAGCGGAG	AGTTTGGCAT	TACAAGATCG	17430
25	TACCATGAAA	ATGTTTTTTA	ACCCAGGTGT	TGCGCCTTTC	ATGCTTCTGC	AGGAGGGCCA	GAGCCACAGC	17500
	TGCATGTTAC	CGCCTTTGCA	CCAGCTCCAG	AGGCTTGGGA	CCAGGCTGTC	TCAGTTCCAG	GSTGCGTCCG	17570
	GTCGCAGACC	CCCTCCCTCT	TGCCCTTCTT	CTCTGCTCTA	AATCTTCCCT	CGTTTGCATC	TCCTTGAGCC	17640
	GTGCTGGGGC	CTCTGTGCAA	CGCTGTGATC	TCCTTTCCGG	AAACCTCTGG	GCGTGTCTGG	ATACAGGTGC	17710
	CTGACGAGGC	TGGAGGTGTC	TGACATCTGT	GTTGACCCCA	GGGTCAGCT	GGCGTCTGTT	GGGCCCTCCT	17780
30	GGGCGATGAT	GAGGTTCAGAG	GAGTTTTCAG	AGGTTAAAAA	TCTGTGGAAA	CTCCAGGGGC	CATGTGACCT	17850
	GCCACCTGAT	CCTCCCATAT	TCAGCTTCAG	CTTGCTCTCA	TTTCCCCACC	AGGGTCTCTA	GCTCCGAGGA	17920
	GTCCTCCGTG	AGGGCCTGGG	CTCAGGCGAG	GGCGGCTGAG	TTTCCCCACC	CATGTGGGGA	CCCTTGGGTA	17990
	GTCTGTTGAT	TGGGTAGCCC	TGAGGAGGCC	GAGATGCGAT	GGGCCACGGG	CGGTTTCCAA	ACACAGAGTC	18060
	AGGCACGTGG	AAGGCCCAGG	AATCCCCCTC	CCTCGAGGCA	GAGATGGGAG	AACGGGAGAG	TGGGCCCTCA	18130
35	TTTACGGCGA	CGCCAGGCTG	AGTGGCGCAG	CTCTGTGTGG	TCACGTGGGC	CTGGGGGGTG	GGGTCTGATT	18200
	CAACATCCGT	CGAGCTCCGG	CTTCTCGGCC	CGTCTGGGCC	CGCCCTCCAC	ACGGGCTAGG	GGTGGACGCC	18270
	CCGACTCTCTA	GAGGTTGGCT	ATTCTTCCCT	TGGAAGAGA	CGCCCTCACC	CATGCTTAGT	GTITCCCTCT	18340
	TGGGTACAGGA	CGGTGGCCGT	GTGGCAACCC	CGGGACCTTA	GGCTTATTTA	TTTGTTTAAA	AACATTCTTG	18410
	CGGTGGCTCT	CGTTTGTGCT	AAATGGGGAA	AAGACATCCC	ACCTCAGCAG	AGGTACTGAG	AGGCTGAAAC	18480
40	CGGGGTGCTG	CGTTGACTGT	TGTGATCTCA	GGCTATTCCA	GAAATGGGCT	AGGAATGCTAG	TGAGACACAG	18550
	TACATGGGGG	GCTCAGGACG	TGGGTGAGAT	GAGTACACGG	GGGGGCTCAG	CGAGTGGGTT	AGGCCAGGTA	18620
	CTAGGGGGGC	TCAGGCATGT	GGTGAGATGA	GGTACACGCG	GGGCTCAGCG	AGAGGGTCTG	ACCCAGGTACA	18690
	CGGGGGCTCT	GATCACACGC	ACATATGAGC	ACATGTGCAC	ATGTGCTGTT	TCATGGTACG	CAGTCTGTGT	18760
	CACACCTGCC	CCAAAGTCCC	AGGAAGTCGA	GAGGCCAAAG	ATGGAGGCTG	ACAGGGCTGG	CGGGTGGGCT	18830
45	CACACCTGTA	GTCCAGACAC	TTTGGGAGGC	CGAGCGGAGA	GGATCCCTTG	AGCCGAGGAG	TTTAAGACCA	18900
	GCCTGAGCAA	CATAGTAGAA	CCCCATCTCT	ATGAAAATA	AAAAACAAAA	TGTAGTGAAC	ATGGTGGTGT	18970
	CGCGCTGTAG	TTCCAAATCT	TGGGAGGCTG	A				

TGCCCCGCCA CCCACACGTC CTAGGAGGGT TGGAGGATGC CACCTCTGGC CTCTTCTGGA ACGGAGTCTG 21280
 ATTTTGGCCCC CGCAGCCAG ACGCAGCTGA GTCCGAAGCT CCCGGGGACG ACGCTGACTG CCCTGGAGGC 21350
 CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT GGCCACCCGC CCACAGCCAG 21420
 GCCGAGAGCA GACACCAGCA GCCCTGTAC GCCGGGCTCT ACGTCCCAGG GAGGGAGGGG CGGCCACAC 21490
 CCAGGCCCGC ACCGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCGAG GCCTGCATGT CCGGCTGAAG 21560
 GCTGAGTGTG CCGCTGAGGC CTGAGCGAGT GTCCAGCCAA GGGCTGAGTG TCCAGCACAC CTGCCGTCTT 21630
 CACTTCCCCA CAGGCTGGCG CTCGGCTCCA CCCCAGGGCC AGCTTTTCTT CACCAGGAGC CCGGCTTCCA 21700
 CTCGCCACAT AGGAATAGTC CATCCCCAGA TTCGCCATTG TTCACCCCTC GCCTTGCCCT CTTTGGCCTT 21770
 CCACCCCCAC CATCCAGGTG GAGACCTGA GAAGGACCTT GGGAGCTCTG GGAATTTGGA GTGACCAAG 21840
 GTGTGCCCTG TACACAGCG AGGACCTGAC ACCTGGATGG GGGTCCCTGT GGGTCAAAAT GGGGGGAGGT 21910
 GCTGTGGGAG TAAAATACTG AATATATGAG TTTTTCAGTT TTGAAAAAAA TCTCATGTTT GAATCCTAAT 21980
 GTGCACTGCA TAGACACCAC TGTATGCAAT TACAGAAGCC TGTGAGTGAA CGGGGTGGTG GTCAGTGGCG 22050
 GCCCATGGCC TGGCTGTGCA TTTACGGAAG TCTATGAGTG AATGGGGTTG TGGTCACTGC GGGCCCATGG 22120
 CCTGGCTGGG CTTGGGAGGT TTCTGATGCT GTGAGGCAGG AGGGGAAGGA GGGTAGGGGA TAGACAGTGG 22190
 GAGCCCCCAC CTTGGAAGAC ATAACAGTAA GTCCAGGCCG GAAGGGCAGC AGGGATGCTG GGGGCCAGC 22260
 TTGGGCGGCG GGGATGATGG AGGGCTGGCC CAGGGTGGCA GGGATGATGG GGGCCCGCAG GGGGGTGG 22330
 GGGGTGATGG GGGGGGCTGG TCTGGGTGGC GGGGAAGATG GGGGAAGCTG GCTGGGCCCT CTCCTCCCTT 22400
 GCCTCCACCC TGCAGCCGTG GATCCGGATG TGCTTCCCTG GTGCACATCC TCTGGGCCAT CAGCTTTCTT 22470
 GGAGGTGGGG GGCAGGGGCA TGACACCATC CTGTATAAAA TCCAGGATTG CTCCTCTCTG ACGCCCCAAC 22540
 TCAGGTTGAA AGTCACATTC CGCCTCTGGC CATCTCTTAA AGAGTAGACC AGGATTCTGA TCTCTGAAGG 22610
 GTGGGTAGGG TGGGGCAGTG GAGGGTGTGG ACACAGGAGG CTTAGGGTGG GGGCTGGTGA TGCTCTCTCA 22680
 TCCTCTTATC ATCTCCAGT CTCATCTCTC ATCTCTTAT CATCTCCAG TCTCATCTGT CTCCTCTTAA 22750
 TCTCCAGTGC TCATCTGTCA TCCTCTTACC ATCTCCAGT CTCATCTCTT ATCTCTTAT CTCCTAGTCT 22820
 CATCCAGACT TACCTCCAG GGGGGGTGCC AGGCTCCGAG TGGAGCTGGA CATACGTCCT TCCTCAGGCA 22890
 GAAAGGAATG GAAGGATTGC AGAGAACAGG AGGGGCGGCT CAGAGGGAGC CAGCTTTGGG GTGAAGAAAC 22960
 AGCCCTCTCT CAGAAGTTGG CTTGGGCCAC ACGAAACCGA GGGCCCTGCG TGAGTGGCTC CAGAGCCTTC 23030
 CAGCAGGTCC CTGTTGGGCG CTTATGGTAT GGCCGGGTCC TACTGAGTGC ACCTTGGACA GGGCTTCTGG 23100
 TTTGAGTGCA GCCCGGACGT GCCTGGTGTG GGGGTGGGGG CTTATGGCCA CTGGATATGG CGTCATTAT 23170
 TGCTGTCTGT TCAGAGAATG TCTGAGTGAC CGAGCCTAAT GTGTATGGTG GGCCCAAGTC CACAGACTGT 23240
 GTCGTAAATG CACTCTGGTG CTTGGAGCCC CCGTATAGGA GCTGTGAGGA AGGAGGGGCT CTTGGCAGCC 23310
 GGCTTGGGGG CGCCTTTGCC CTGCAAACTG GAAGGGAGCG GCCCGGGGCG CCGTGGGCGG ACGACCTCAA 23380
 GTGAGAGGTT GGACAGAAAC GGGCGGGGAC TTCCAGGAG CAGAGGCCGC TGCTCAGGCA CACCTGGGTT 23450
 TGAATCAGAG ACCAACAAGT CAGGCCATTG TTCAGCTATC CATCTCTAC AAAGCTCCAG ATTCCTGTTT 23520
 CTCGGGGTGT TTTTGTGTA AATTTTACTC AGGATTACTT ATATTTTGT CTAAAGTATT AGACCCCTAA 23590
 AAAAGGTATT TGCTTTGATA TGGCTTAACT CACTAAGCAC CTACTTTATT TGCTCTGTTT TATTTATTAT 23660
 TATTATTATT ATTAGAGATG GTGTCTACTC TGTCACCCAG GTTGTAGTG CAGTGGACA GTCATGGCTC 23730
 GCTGTAGCCG CAAACCCCA GGCTCAAATG ATCTCCGCG CTCAGCTTCC CAGAGTGCTG GGATTACAG 23800
 TGTGAGCCAC TGCCCTTGCC TGGCACTTTT AAAAACCACT ATGTAAGGTC AGGTCCAGT GCTTCCACAC 23870
 CTGTATCCC AGTAGTTTGG GAAGCCGAGG CAGAAGGATT GTCTGAGGCC AGGAGTTTGA GACCAGCATG 23940
 GGTAAACATG GGAGACCCCA TCTCTACAAA AAATGCAAAA AGTTATCCGG GCGTGGGCTC CAGCATCTGT 24010
 AGTCCCAGCT GCTCGGGAGG CTGAGTGGGA GGATCGCTTG AGCCCGGGAG GTCATGGCTG CAGTGAGCTG 24080
 TGATTGTACC ATCGCACTCC AGCCTGGGCA ACAGAGTGAG ACCCTGTCTC AAAAAAAGG AAAAAAAG 24150
 AAGGAGAAGG AGAAGAGAAG AAGAAGGAAG AAGGAAGAG AAGAAGAGG AAGAAGGAAG AAGAAGGAG 24220
 AAGGAGGCTT GCTAGGTGCT AGGTGAGACT TCAATCTCA GAGCAAAATG AAATATACAA AGTTTAAAG 24290
 GGAAAGAAAA ACCCAGCTC TTTGGACTTC CTTAGGCCCTG AACTTCACTC CAAGCAGCTT CTTCCACAG 24360
 ACAAGCGTGT ATGGAGCGAG TGAGTTCAAA CGAGAAAGGG AGGAGAAGCA CGCAAGGGTG GAGGCTGTGG 24430
 GTGACACCAG CCAGGACCCC TGAAGGGGAG TGGTTGTTTT CTTGCTCAG CCCACGCTC CTGCCGGTCC 24500
 TGCACCTGCT GTAACCGTCG ATGTTGGTGC CAGGTGCCCA CTTGGGAAGG ATGCTGTGCA GGGGGCTTGC 24570
 CAAACTTTGG TGGGTTTCAG AAGCCCAAGG CACTTGTGGC AGGCACAAAT ACAGCCCTC CCAAGAGATG 24640
 CCCACGTCCT TCTCTGGAA CCTGTGAATG TGTACCCCGC AAGGCAGAGG CTGGTGAAGG CTGCAGGTGG 24710
 AATCACGGCT GCCAGTCAGC CGATCTTAAG GTCATCTGG ATTATCTGGT GGGCCTGATA TGGCCACAG 24780
 GGTCCCTAGA AGTGAGAGAG GGAGGCAGGG GAGACTCAGA GAGGGGACGT GAGAAGGACC ACTGCCACT 24850
 GCTGGCTTTG AGATGGAGGA GGGGGTCCCC AGCCAAGGAA TGGGGGACG CGCTCCATGC TGGAAAAGCA 24920
 AGCAATCCTC CCCGGTCTG AGGGCACACG GCCCTGCCCA CGCCTCGATT TCAGGCCAGT GGGACCTGTT 24990
 TCAGCTTTCC GGCCCTCAGA GCTGTAAGAT GATCGGTTTG TGTTCAGCCA CTAAGCTGCA GTGATTCGTC 25060
 ACAGCAGCAA ATGGAATAGC AGTACAGGGA AATGAATACA GGGACAGTTC TCAGAGTGAC TCTCAGCCCA 25130
 CCCCTGGG 25138

60 Beispiel 5

Der Vergleich der oben beschriebenen genomischen hTC-Sequenz mit der Sequenz
 der hTC-cDNA (Fig. 6; entsprechend SEQ ID NO 2) ermöglichte die Aufklärung der
 Exon-Intron-Struktur des hTC-Gens. Die genomische Organisation des hTC-Gens ist
 in Fig. 7 schematisch dargestellt. Die kodierende Region des hTC-Gens setzt sich aus

16 Exons zusammen, die in ihrer Größe zwischen 62 bp und 1354 bp variieren (s. Tabelle 1). Exon 1 enthält das Translationsstartcodon ATG. Das Translationsstopcodon TGA sowie der 3' untranslatierte Bereich liegen auf Exon 16 (Fig. 8). Ein mögliches Polyadenylierungssignal (AATAAA) wurde weder in Exon 16 noch in
 5 den 3195 bp der folgenden 3'-flankierenden Region gefunden. Basierend auf der Konsensussequenz

	5'-Exon			Intron				3'-Exon			
Prä-mRNA	A/C	A	G		G	T	A/G	A	...	N C	A G G
10 Häufigk.(%)	70	60	80		100	100	95	70		80	100 100 60

wurden die Exon-Intron-Übergänge bestimmt und in Tabelle 1 aufgeführt. Mit Ausnahme der 5'-Splice-Stelle zwischen Exon 15 und Intron 15 stimmen alle Exon-Intron-Übergänge mit der publizierten (Shapiro und Senapathy, 1987) Splice-Konsensussequenz überein. Die Größe der Introns liegt zwischen 104 bp und 8616
 15 bp. Da Intron 6 nur zum Teil isoliert wurde, kann die exakte Länge des hTC-Gens nicht bestimmt werden. Basierend auf der von Intron 6 erhaltenen Teilsequenz von ~4660 bp beträgt die minimale Größe des hTERT Gens 37 kb.

Die Introns 1-5 sowie der 5'-Bereich des Introns 6 sind in Contig 1 enthalten:

Intron 1: bp 11493-11596 (SEQ ID NO 4);

Intron 2: bp 12951-21566 (SEQ ID NO 5);

Intron 3: bp 21763-23851 (SEQ ID NO 6);

5 Intron 4: bp 24033-24719 (SEQ ID NO 7);

Intron 5: bp 24900-25393 (SEQ ID NO 8);

5'-Bereich von Intron 6: bp 25550-26414 (SEQ ID NO 9).

Der 3'-Bereich des Introns 6 sowie die Introns 7-15 sind in Contig 2 an folgenden
10 Positionen lokalisiert:

3'-Bereich von Intron 6: bp 1-3782 (SEQ ID NO 10);

Intron 7: bp 3879-4858 (SEQ ID NO 11);

Intron 8: bp 4945-7429 (SEQ ID NO 12);

Intron 9: bp 7544-9527 (SEQ ID NO 13);

15 Intron 10: bp 9600-11470 (SEQ ID NO 14);

Intron 11: bp 11660-15460 (SEQ ID NO 15);

Intron 12: bp 15588-16467 (SEQ ID NO 16);

Intron 13: bp 16530-19715 (SEQ ID NO 17);

Intron 14: 19841-20621 (SEQ ID NO 18);

20 Intron 15: 20760-21295 (SEQ ID NO 19).

Der 3'-nichttranskribierte Bereich befindet sich ebenfalls im Contig 2 an Position
21960-25138 (SEQ ID NO 20).

25 Die genannten Introns haben im einzelnen folgende Sequenzen:

Intron 1 (SEQ ID NO 4)

GTGGGCCTCCCCGGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGG
CGACTCAGGGCGCTTCCCCCGCAG

5

Intron 2 (SEQ ID NO 5)

GTGAGGAGGTGGTGGCCGTCGAGGGCCAGGCCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCC
CTGGTCCTCCTGTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTCGAGTGGACACGGTGATCTCTGCC
TCTGTCTCCCTCCTGTCCAGTTTGCATAAACTTACGAGGTTACCTTCACGTTTGTATGGACACGCGTTTCCAGGCGC
10 CGAGGCCAGAGCAGTGAACAGAGGAGGCTGGGCGCGGCAGTGGAGCCGGGTTGCCGCAATGGGGAGAAAGTGTCTGGAAG
CACAGACGCTCTGGCGAGGGTGCTGCAGGTTACCTATAATCCTCTTCGCAATTTCAAGGGTGGGAATGAGAGGTGGGGA
CGAGAACCCCTCTTCTGGGGGTGGGAGGTAAGGGTTTTCAGGTGCACGTGGTCAGCCAATATGCAGGTTTGTGTTTA
AGATTTAATTGTGTGTTGACGGCCAGGTGCGGTGGCTCACGCCGGTAATCCAGCACTTTGGGAAGCTGAGGCAGGTGGA
TCACCTGAGGTGAGGAGTTTGAGACCAGCTGACCAACATGGTGAACCCCTATCTGTACTAAAAATACAAAAATTAGCTG
15 GGCATGGTGGTGTGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTGAACCCAGGAGGCGGAGGC
TGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCGACAAGAGTGAACTCTGTCTTAAAAAAAAGTGT
CGTTGATTGTGCCAGGACAGGGTAGAGGGAGGAGATAAGACTGTTCTCCAGCACAGATCCTGGTCCCCTCTTTAGGTAT
GAAGAGGGCCACATGGGAGCAGAGGACAGCAGATGGCTCCACCTGCTGAGGAAGGGACAGTGTGTTGGTGGTGTTCAGGGG
ATGGTGTCTGTGGGCCCTGCCGTGTCCCAACCTGTTTTCTGGATTGTATGTTGAGGAACCTCCGCTCCAGCCCCCTTT
20 TGGCTCCAGTGCTCCAGGCCCTACCGTGGCAGCTAGAAGAAGTCCCGATTTACCCCCCTCCCAAACTCCCAAGAC
ATGTAAGACTTCCGGCCATGCAGACAAGGAGGGTGACCTTCTGGGGCTCTTTTTTTCTTTTTTTCTTTTATGGTGGC
AAAAGTCATATAACATGAGATTGGCACTCCTAACACCGTTTTCTGTGTACAGTGCAGAATTGCTAACTCGGCGGTGTTTA
CAGCAGGTTGCTTGAAATGCTGCGTCTTGCCTGACTGGAAGTCCCTACCCATCGAACGGCAGCTGCCTCACACCTGTGTC
GGCTCAGGTGGACCACGCCAGTCAGATAAGCGTCATGCAACCCAGTTTTGCTTTTTGTGCTCCAGCTTCCTTCGTTGAG
25 GAGAGTTTGAGTTCTCTGATCAGGACTCTGCCTGTCAATGCTGTTCTCTGACTTCAGATGAGGTCAAACTGCCCCCTGG
CTTATGCAGGGAGTGAGGCGTGGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGGCGTTGCCCCCAGGTGTCCCT
GTACAGTGTAGGGTGAGTGAGGCGCGGCCCGGGTGTCCCTGTCCCGTGCAGCGTGATTGAGGTGTGGCCCCCGGGTGT
CCCTGTACAGTGTAGGGTGAGTGAGGCGCCATCCCCGGGTGTCCCTGTACGTGTAGGGTGAGTGAGGCGTGGTCCCCGG
GTGTCCCTGTCCCGTGCAGGGTGAGTGAGGCACTGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGGCGCGGTCC
30 CCGGGTGTCCCTCTCAGGTGTAGGGTGAGTGAGGCGCGGCCCCAGGGTGTCCCTGTACGTGTAGGGTGAGTGAGGCACC
GTCCCTGGGTGTCCCTCCAGGTATAGGGTGAGTGAGGCACTGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGG
CGCGGCCCCCGGGTGTCCCTCTCAGGTGCAGGGTGAGTGAGGCGCTGTCCCTGGGTGTCCCTGTCTCGTGTAGGGTGAGT
GAGGCTCTGTCCCAAGGTGTCTTGGCGTTTGCTCACTTGAGCTTGCTCCTGAATGTTTGTCTTTCTATAGCCACAGCT
GCGCCGGTTGCCATTGCCTGGGTAGATGGTGAGGCGCAGTGCTGGTCCCCAAGCCTATCTTTCTGATGCTCGGCTCT
35 TCTTGGTCACCTCTCCGTTCCATTTTGCTACGGGGACACGGGACTGCAGGCTCTCGCTCCCGGTGCCAGGCACTGCAG
CCACAGCTTCAGGTCGCTTGCCTCTGTGGGCTGGCTTGCTACCAAGTGCCCGCCACATGCATGCTGCCAATACTCC
TCTCCAGCTTGTCTCATGCCAGGCTGGACTCTGGGCTGCCTGTGTCTGTGCCACGTGTTGCTGGAGACATCCAGAA
AGGGTTCTCTGTGCCCTGAAGGAAAGCAAGTCACCCAGCCCCCTCACTTGTCTGTTTTCTCCCAAGCTGCCCTCTGC
TTGGCCCCCTTGGGTGGGTGGCAACGCTTGTACCTTATTCTGGGCACCTGCCGCTCATTGCTTAGGCTGGGCTCTGCCT
40 CCAGTCGCCCCCTCACATGGATTGACGTCCAGCCACAGGTTGGAGTGTCTCTGTCTGTCTCTGCTCTGAGACCCACGTG

GAGGGCCGGTGTCTCCGCCAGCCTTCGTGAGACTTCCCTCTGGGTCTTAGTTTGAATTTCACTGATTTACCTCTGACG
TTTCTATCTCTCCATTGTATGCTTTTCTGGTTTATTCTTTTCATTCTTTCTAGCTTCTTAGTTTAGTCATGCCTTTC
CCTCTAAGTGCTGCCTTACCTGCACCCCTGTGTTTTGATGTGAAGTAATCTCAACATCAGCCACTTTCAAGTGTCTTAAA
ATACTTCAAAGTGTAAATACTTCTTTTAAGTATTCTTATTCTGTGATTTTTTCTTTGTGACGCTGTGTTTTGACGTGA
5 AATCATTTTGATATCAGTGACTTTTAAGTATTCTTAGCTTATTCTGTGATTTCTTTGAGCAGTGAGTTATTTGAACACT
GTTTATGTTCAAGATATGTAGAGTATCAAGATACGTAGAGTATTTAAGTTATCATTTTATTATTGATTTCTAACTCAGT
TGTTAGTGGTCTGTATAATACCAATTATTTGAAGTTTGGGAGCCTTGCTTTGTGATCTAGTGTGTGCATGGTTTCCAG
AACTGTCCATTGTAAATTTGACATCCTGTCAATAGTGGGCATGCATGTTCACTATATCCAGCTTTATTAAGGTCCAGTGCA
AAGCTTCTGTCTCCTTCTAGATGCATGAAATTTCAAGAAGGAGGCCATAGTCCCTCACCTGGGGGATGGGTCTGTTTATT
10 TCTTCTCGTTTGGTAGCATTTATGTGAGGCATTGTTAGGTGCATGCACGTGGTAGAATTTTATCTTCTGATGAGTGAA
TCTTTTGGAGACTTCTATGTCTCTAGTAATCTAGTAATCTTTTTTAAATTGCTCTTAGTACTGCCACACTGGGCTTCT
TTTGATTAGTATTTTCTGCTGTGCTGTTTTCTGCCTTTAAATTTATATATATATATATATTTTTTTTTTTTTTGGAGACA
GAGTCTTGGTCTGTGCGCCAGGGTGTGAGTGCAGTGGTGTGATCACAGGTCAGTGTAACCTTTTACCTTCTGGCCTGAGCCGT
CCTCTCACCTCAGCCTCCTGAGTAGCTGGAACGTGCAGACACGACCGCTACACCTGGCTAATTTTTAAATTTTTCTGGA
15 GACAGGGTCTTGCTGTGTTGCCAGGCTGGTCTCAAACCTCTGGACTCAAGGGATCCATCTACCTCGGCTTCCCAAAGTG
CTGAATTACAGGCATGAGCCACCATGTCTGGCTAATTTTCAACACTTTTATATTCTTATAGTGTGGGTATGCTCTGTTA
ACAGCATGTAGGTGAATTTCAATCCAGTCTGACAGTCGTGTTTAACTGGATAACCTGATTATTTTCACTTTTTTGTGTC
ACTAGAGACCCGCTGGTGACTCTGATTCTCCACTTGCCTGTTGATGTCCTCGTTCCCTGTTTCTCACCACCTCTTG
GGTTGCCATGTGCGTTTTCTGCGAGTGTGTGATCCTCTCGTTGCCCTCGGTCACTGGGCATTTGCTTTTATTCT
20 CTTTGCTTAGTGTTACCCCTGATCTTTTATTGTGCTGTTGCTTTTGTATTATTGAGACAGTCTCACTCTGTCAACCA
GGCTGGAGTGAATGGCACAATCTCGGCTCACTGCAACCTCTGCCTCCTCGGTTCAAGCAGTTCTCATTCTCAACCTCA
TGAGTAGCTGGGATTACAGGCGCCACCACCGCCTGGCTAATTTTTGTATTTTAGTAGAGATAGGCTTTCACCATGT
TGGCCAGGCTGGTCTCAAACCTCTGACCTCAAGTGATCTGCGCGCTTGGCCTCCACAGTGCTGGGATTACAGGTGCAA
GCCACCGTGCCCGCATACCTTGATCTTTTAAATGAAGTCTGAAACATTGTACCTTGTCTGAGCAATAAGACCCCTT
25 AGTGTATTTTAGCTCTGGCCACCCCCAGCCTGTGTGCTGTTTTCCCTGCTGACTTAGTTCTATCTCAGGCATCTTGACA
CCCCACAAGCTAAGCATTATTAATATTGTTTTCCGTGTTGAGTGTCTGTAGCTTTGCCCCCGCCTGCTTTCTCTCC
TTTGTTCCTGCTGTCTTCTGTCTCAGGCCGCGCTGCGGGTCCCTTCTTGTCTTTGCGTGGTCTTCTGTCTGTG
TTATTGCTGGTAACCCAGCTTTACCTGTGCTGGCCTCCATGGCATCTAGCGACGTCCGGGACCTCTGCTTATGATGC
ACAGATGAAGATGTGGAGACTCACGAGGAGGGCGGTATCTGGCCCGTGAGTGTCTGGAGCACCACGTGGCCAGCGTTC
30 CTTAGCCAGTGAGTGACAGCAACGTCCGCTCGGCTGGGTTGAGCCTGGAAAACCCAGGCATGTGCGGGTCTGGTGGCT
CCGCGGTGTGAGTTTGAAATCGCGCAACCTGCGGTGTGGCGCCAGCTCTGACGGTGTGCTGGCGGGGAGTGTCTG
CTTCTCTCTTCTGCTTGGGAACAGGACAAAGGATGAGGCTCCGAGCCGTTGTGCGCCAAACAGGAGCATGACGTGAGCC
ATGTGGATAATTTTAAATTTCTAGGCTGGGCGCGGTGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCCAAGGCGGG
TGGATCACGAGGTGAGGAGGTGAGACCATCTGGCCCAACATGATGAAACCCCATCTGTACTAAAAACACAAAAATTAGC
35 TGGGCGTGGTGGCGGGTGCTGTAATCCAGCTACTCGGGAGGCTGAGGCAGGAGAAATTGCTTGAACTGGGAGTTGGAA
GTTGCAGTGAGCCGACATTGCACCACTGCCTCCAGCCTGGCAACACAGCGAGACTCTGTCTCAAAAAAAAAAAAAA
AAAAAAAAAAATTTAGTAGCCACATTAAAAAAGTAAAAAAGAAAAGTGAAATTAATGTAATAATAGATTTTACTGAA
GCCCAGCATGTCCACACCTCATTTTTAGGGTGTATTGGTGGGAGCATCACTCACAGGACATTGACATTTTTTGAGC
TTTGTCTGCGGGATCCCGTGTGTAGGTCCCGTGTGCGTGGCCATCTCGGCCTGGACCTGTGGGCTTCCCATGGCCATGGCT
40 GTTGTACCAGATGGTGAGGTCCGGATGAGGTGCGCAGGCCCTCAGTGAGCTGGATGTGCAGTGTCCGGATGGTGACG
TCTGGGATGAGGTGCGCAGGCCCTGTGTGAGCTGGATGTGTGGTGTCTGGATGGTGAGGTGAGGGTGTGAGGTCTCCAG

5 G C C C T C G G T G A G C T G G A G G T A T G G A G T C C G G A T G A T G C A G G T C C G G G G T G A G G T C G C C A G G C C C T G C T G T G A G C T G G A T G
T G T G G T G T C T G G A T G G T G C A G G T C A G G G G T G A G G T C T C C A G G C C C T C G G T A A G C T G G A G G T A T G G A G T C C G G A T G A T G C A
G G T C C G G G G T G A G G T C G C C A G G C C C T G C T G T G A G C T G G A T G T G T G G T G T C T G G A T G G T G C A G G T C T G G G G T G A G G T C A C C
A G G C C C T G C G G T G A G C T G G G T G T G C G G T G T C T G G A T G G T G C A G G T C T G G A G T G A G G T C G C C A G A C G G T G C C A G A C C A T G C
10 G G T G A G C T G G A T A T G C G G T G T C C G G A T G G T G C A G G T C T G G G G T G A G G T T G C C A G G C C C T G C T G T G A G T T G G A T G T G G G G T
G T C C G G A T G C T G C A G G T C C G G T G T G A G G T C A C C A G G C C C T G C T G T G A G C T G G A T G T G T G G T G T C T G G A T G G T G C A G G T C T
G G G G T G A A G G T C G C C A G G C C C T G C T T G T G A G C T G G A T G T G T G G T G T C T G G A T G G T G C A G G T C T G G A G T G A G G T C G C C A G
G C C C T C G G T G A G C T G G A T G T G C A G T G T C C A G A T G G T G C A G G T C C G G G G T G A G G T C G C C A G A C C C T G C G G T G A G C T G G A T G
T G C G G T G T C T G G A T G G T G C A G G T C T G G A G T G A G G T C G C C A G G C C C T C G G T G A G C T G G A T G T A T G G A G T C C G G A T G G T G C C
15 G G T C C G G G G T G A G G T C G C C A G A C C C T G C T G T G A G C T G G A T G T G C G G T G T C T G G A T G G T A C A G G T C T G G A G T G A G G T C G C C
A G A C C C T G C T G T G A G C T G G A T A T G C G G T G T C C G G A T G G T G C A G G T C A G G G G T G A G G T C T C C A G G C C C T C G G T G A G C T G G A
G G T A T G G A G T C C G G A T G A T G C A G G T C C G G G G T G A G G T C G C C A G G C C C T G C T G T G A A C T G G A T G T G C G G C G T C T G G A T G G T
G C A G G T C T G G G G T G T G G T C G C C A G G C C C T C G G T G A G C T G G A G G T A T G G A G T C C G G A T G A T G C A G G T C C G G G G T G A G G T C G
C C A G G C C C T G C T G T G A G C T G G A T G T G C G G C G T C T G G A T G G T G C A G G T C T G G G G T G T G G T C G C C A G G C C C T C G G T G A G C T G
20 G A G G T A T G G A G T C C G G A T G A T G C A G G T C C G G G G T G A G G T T G C C A G G C C C T G C T G T G A G C T G G A T G T G C T G T A T C C G G A T G
G T G C A G T C C G G G G T G A G G T C G C C A G G C C C T G C T G T G A G C T G G A T G T G C T G T A T C C G G A T G G T G C A G G T C T G G G G T G A G G T
C A C C A G G C C C T G C G G T G A G C T G G T T G T G C G G T G T C C G G T G T G C A G G T C C G G G G T G A G T T C G C C A G G C C C T C G G T G A G C
T G G A T G T G C G G T G T C C C C G T G T C C G G A T G G T G C A G G T C C A G G G T G A G G T C G C T A G G C C C T T G G T G G G C T G G A T G T G C C G T
G T C C G G A T G G T G C A G G T C T G G G G T G A G G T C G C C A G G C C T T T G G T G A G C T G G A T G T G C G G T G T C T G C A T G G T G C A G G T C T G
25 G G G T G A G G T C G C C A G G C C C T T G G T G G G C T G G A T G T G T G G T G T C C G G A T G G T G C A G G T C C G G C G T G A G G T C G C C A G G C C C T
G C T G T G A G C T G G A T G T G C G G T G T C T G G A T G G T G C A G G T C C G G G G T G A G G T A G C C A A G G C C T T C G G T G A G C T G G A T G T G G G
G T G T C C G G A T G G T G C A G G T C C G G G G T G A G G T C G C C A G G C C C T G C G G T T A G C T G G A T A T G C G G T G T C C G G A T G G T G C A G G T
C C G G G G T G A G G T C A C C A G G C C C T G C G G T T A G C T G G A T G T G C G G T G T C T G G A T G G T G C A G G T C C G G G G T G A G G T C G C C A G G
C C C T G C T G T G A G C T G G A T G T G C T G T A T C C G G A T G G T G C A G G T C C G G G G T G A G G T C G C C A G G C C C T G C A G T G A G C T G G A T G
30 T G C T G T A T C C G G A T G G T G C A G G T C T G G C G T G A G G T C G C C A G G C C C T G C G G T T A G C T G G A T A T G C G G T G T C G G A T G G T G C A
G G T C C G G G G T G A G G T C A C C A G G C C C T G C G G T T A G C T G G A T G T G C G G T G T C C G G A T G G T G C A G G T C T G G G G T G A G G T C G C C
A G G C C C T G C T G T G A G C T G G A T G T G C T G T A T C C G G A T G G T G C A G G T C C G G G G T G A G G T C G C C A G G C C C T G C G G T G A G C T G G
A T G T G C T G T A T C C G G A T G G T G C A G G T C T G G C G T G A G G T C G C C A G G C C C T G C G G T G A G C T G G A T G T G C A G T G T A C G G A T G G
T G C A G G T C C G G G G T G A G G T C G C C A G G C C C T G C G G T G G G C T G T A T G T G T G T G T C T G G A T G G T G C A G G T C C G G G G T G A G T T
35 C G C C A G G C C C T G C G G T G A G C T G G A T G T G T G G T G T C T G G A T G C T G C A G G T C C G G G G T G A G T T C G C C A G G C C C T C G G T G A G C
T G G A T A T G C G G T G T C C C C G T G T C C G A A T G G T G C A G G T C C A G G G T G A G G T C G C C A G G C C C T T G G T G G G C T G G A T G T G C C G T
G T C C G G A T G G T G C A G G T C T G G G G T G A G G T C G C C A G G C C C T T G G T G A G C T G G A T G T G C G G T G T C C G G A T G G T G C A G G T C C G
G G G T G A G G T C A C C A G G C C C T C G G T G A T C T G G A T G T G G C A T G T C C T T C T G T T T A A G

35 Intron 3 (SEQ ID NO 6)

G T A C T G T A T C C C C A C G C C A G G C C T C T G C T T C T C G A A G T C C T G G A A C A C C A G C C C G G C C T C A G C A T G C G C C T G T C T C C A C T
T G C C T G T G C T T C C C T G G C T G T G C A G C T C T G G G C T G G G A G C C A G G G G C C C C G T C A C A G G C C T G G T C C A A G T G G A T T C T G T G
C A A G G C T C T G A C T G C C T G A G C T C A C G T T C T C T T A C T T G T A A A A T C A G G A G T T T G T G C C A A G T G G T C T C T A G G G T T T G T A
A A G C A G A A G G G A T T T A A A T T A G A T G G A A C A C T A C C A C T A G C C C T C C T T G C C T T C C C T G G G A T G T G G G T C T G A T T C T C T C
40 T C T C T T T T T T T T C T T T T T G A G A T G G A G T C T C A C T C T G T T G C C A G G C T G G A G T G C A G T G G C A T A A T C T T G G C T C A C T

5 GCAACCTCCACCTCCTGGGTTTAAAGCGATTACACAGCCTCAGCCTCCTAAGTAGCTGGGATTACAGGCACCTGCCACCAC
GCCTGGCTAATTTTGTACTTTTAGGAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCATGACCTCAGG
TGATCCACCCACCTTGGCCTCCCAAAGTGCTGGGTTTACAGGCTAAGCCACCGTGCCAGCCCCGATTCTCTTTAATT
CATGCTGTTCTGTATGAATCTTCAATCTATTGGATTAGGTATGAGAGGATAAAATCCACCCACTTGGCGACTCACTG
10 CAGGGAGCACCTGTGCAGGGAGCACCTGGGGATAGGAGAGTTCCACCATGAGCTAACTTCTAGGTGGCTGCATTGAATG
GCTGTGAGATTTTGTCTGCAATGTTGCGCTGATGAGAGTGTGAGATTGTGACAGATTCAAGCTGGATTGCAATCAGTGAG
GGACGGGAGCGCTGGTCTGGGAGATGCCAGCCTGGCTGAGCCCAGGCCATGGTATTAGCTTCTCCGTGTCCCGCCAGGC
TGACTGTGGAGGGCTTTAGTCAGAAGATCAGGGCTTCCCAGCTCCCTGCACACTCGAGTCCCTGGGGGGCCTTGTGAC
ACCCCATGCCCAAATCAGGATGTCTGCAGAGGGAGCTGGCAGCAGACCTCGTCAGAGGTAACACAGCCTCTGGGCTGGG
15 GACCCCGACGTGGTGTGGGGCCATTTCTTGCATCTGGGGAGGGTCAGGGCTTCCCTGTGGGAACAAGTTAATACAC
AATGCACCTTACTTAGACTTTACACGTATTTAATGGTGTGCGACCCAACATGGTCATTTGACCAGTATTTTGGAAAGAAT
TTAATTGGGGTGACCGGAAGGAGCAGACAGACGTGGTGGTCCCAAGATGCTCCTTGTCACTACTGGGACTGTTGTTCTG
CCTGGGGGGCCTTGGAGGCCCTCTCCCTGGACAGGGTACCGTGCCCTTTCTACTCTGCTGGGCCTGCGGCCTGCGGT
AGGGCACCAGCTCCGGAGCACCCGCGGCCCCAGTGTCACGGAGTGCCAGGCTGTGAGCCACAGATGCCAGGTCCAGGT
20 GTGGCCGCTCCAGCCCCCGTGCCCCATGGGTGGTTTGGGGGAAAAGGCCAAGGGCAGAGGTGTGAGGAGCTGGTGGG
CTCATGAGAGCTGATTCTGCTCCTTGGCTGAGCTGCCCTGAGCAGCCTCTCCCGCCCTCTCCATCTGAAGGATGTGGCT
CTTTCTACCTGGGGTCTGCTGGGGCCAGCCTTGGGTACCCCAAGTGGCTGTACCAGAGGGACAGGCATCCTGTGTGG
AGGGGCATGGGTTACGTGGCCCCAGATGCAGCCTGGGACCAGGCTCCCTGGTGTGATGGTGGGACAGTACCCTGGGG
GTTGACCGCGGACTGGGCGTCCCAGGGTTGACTATAGGACCAGGTGTCCAGGTGCCCTGCAAGTAGAGGGGCTCTCAG
25 AGGCGTCTGGCTGGCATGGGTGGACGTGGCCCGGGCATGGCCTTCAGCGTGTGCTGCGTGGGTGCCCTGAGCCCTCAC
TGAGTCGGTGGGGCTTGTGGCTTCCGTGAGCTTCCCCCTAGTCTGTTGTCTGGCTGAGCAAGCCTCCTGAGGGGCTCT
CTATTGCAG

Intron 4 (SEQ ID NO 7)

25 GTGGCTGTGCTTTGGTTTAACTTCCTTTTAAACAGAAGTGCCTTTGAGCCCCACATTTGGTATCAGCTTAGATGAAGGG
CCCGGAGGAGGGGCCACGGGACACAGCCAGGGCCATGGCACGGCGCCAACCCATTTGTGCGCACAGTAGGTTGGCCGAGG
TGCCGGTGCTCCAGAAAAGCAGCGTGGGGGTGTAGGGGAGCTCCTGGGGCAGGGACAGGCTCTGAGGACCACAAGAAG
CAGCCGGGCCAGGGCCTGGATGCAGCACGGCCGAGGTCTGGATCCGTGTCCTGCTGTGGTGCGCAGCCTCCGTGCGCT
30 TCCGCTTACGGGGCCCGGGGACCAGGCCAGACTGCCAGGAGCCACCGGGCTCTGAGGATCTGGACCTTGCCCCACGG
CTCTGCACCCACCCCTGTGGCTGCGGTGGCTGCGGTGACCCCGTCATCTGAGGAGAGTGTGGGGTGAAGTGGACAGAG
GTGTGGCATGAGGATCCCGTGTGCAACACACATGCGGCCAGGAACCCGTTTCAAACAGGGTCTGAGGAAGCTGGGAGGGG
TTCTAGGTCCCGGTCTGGGTGGCTGGGGACACTGGGGAGGGGCTGCTTCTCCCCTGGGTCCCTATGTGGGGTGGGCAG
TTGGCCGGATCCACTTTCCTGACTGTCTCCATGCTGTCCCCGCCAG

35 Intron 5 (SEQ ID NO 8)

GTGGGTGCCGGGACCCCCGTGAGCAGCCCTGCTGGACCTTGGGAGTGGCTGCCTGATTGGCACCTCATGTTGGGTGGAG
GAGGTACTCCTGGGTGGGCCGAGGGAGTGACGGTACCCCTGTCACTGTTGAGGACACACCTGGCACCTAGGGTGGAGGC
CTTCAGCCTTTCTGCAGCACATGGGGCCGACTGTGCACCTGACTGCCCGGGCTCCTATTCCCAAGGAGGGTCCCACTG
GATTCCAGTTTCCGTGAGAGAAGGAACCGCAACGGCTCAGCCACCAGGCCCCGGTGCTTGCACCCAGTCTGTAGCCAG
40 GGGTCTCCTGCTCCTGAGGCTCAGAGAGGGGACACAGCCCGCCTGCCCTTGGGGTCTGGAGTGGTGGGGTTCAGAGAGAG

AGTGGGGGACACCGCCAGGCCAGGCCCTGAGGGCAGAGGTGATGTCTGAGTTTCTGCGTGGCCACTGTCACTCTCCTCGC
CTCCACTCACACAG

5'-Bereich Intron 6 (SEQ ID NO 9)

5 GTAAGGTTACAGTGTGATAGTCGTGCCAGGATGTGTGTCTCTGGGATATGAATGTGTCTAGAATGCAGTCGTGTCTGTG
ATGCGTTTCTGTGGTGGAGGTACTTCCATGATTACACATCTGTGATATGCGTGTGTGGCAGTCGTGTGTGTGGTGCAT
GTATCTGTGGCGTGCATATTTGTGGTGTGTGTGTGTGGCAGTCGTGTGTCCATGGTGTGTGTGCCTGTGGTGTGCATG
TGTGTGTGTCTGTGACAGTCGATGTTTCATGCTGTGTGTGTGCATGTCTGTGATGTGCCTATTTGTGGTGTGTGTGTGCAT
GTGTCCGTGACATATGCGTGTCTATGGCATGGGTGTGTGTGGCCCTTGGCCTTACTCCTTCTCCTCCAGGCATGGTCC
10 GCACCATGTCTCTCAGCTCTCGGGTGTGTGGTTTGGGGAGCTCCACATTGAGGGTCTCACTTCTAGCATGGGTGCCCT
GTCTGTGCAGGGCTGGGCCTTGGAGACTGTAAGCCAGGTTTGGAGGAGAGTAGGGATGTGGTGGTACCTTCTCTGGA
CCCCCTGGCAGCCCCAGGACCCAGTCGTGGCCTATGCCGGCTCCATGAGATATAGGAAGGCTGATTGAGGCCTCGCTCCCC
GGGACACACTCCTCCAGAGCGGCCGGGGCTTGGGGCTCGGCAGGGGTGAAGGGGCCCTGGGCTTGGGTTCCACCC
AGTGGTCATGAGCACGCTGGAGGGTAAGCCCTCAAAGTCGTGCCAGGCCGGGTGCAGAGGTGAAGAAGTATCCCTGGA
15 GCTTCGGTCTGGGGAGAGGCACATGTGAAACCCACAGGACCTCTTCTCTGACTTCTTGAGCT

3'-Bereich Intron 6 (SEQ ID NO 10)

TGTGGGATTGGTTTTTCATGTGTGGATAGGTGGGATCTGTGGGATTGGTTTTTATGAGTGGGGTAACACAGAGTTCAAG
GCGAGCTTTCTTCTGTAGTGGTCTGCAGGTGCTCCAACAGCTTTATTGAGGAGACCATATCTTCTTTGAACTATGGT
20 CGGGTTTATAGTAAGTCAGGGGTGTGGAGGCTCCCTGGGCTCCCTGTTCTGTTTCTTCCACTCTGGGGTGTGTGGT
CCTGCTGTGGTGTGTGGCCGTGGGCAGGGCTTCCAGGCTCCTTGTGTTTATTGGCCTGGATGTGGCCCTGGCTACGCT
CCGTCCTTGAATTTCCCTGCGAGTTGGAGGCTTCTTCTTCTTTTTTCTTTCTTTTTTTTTTTTGTATAACAGA
GTCTCGCTCTTTTTTGCCAGGCTGGAGTGGTTTGGCGTGATCTTGGCTCACTGCAACCTGTGCTTCTGAGTTCAAGCA
ATTCTCTTGCCTCAGCCTCCCAAGTAGCTGGAATTATAGGCGCCACCACCATGCTGACTAATTTTGTAAATTTTAGTAG
25 AGACGAGGTTTCTCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCTCCACCTCGGCCTCCCAAAGT
GCTGGGATGACAGGTGTGAACCGCCGCCCGCCGAGACTCGCTTCTGTCAGCTTCCGTGAGATCTGCAGCGATAGCTG
CCTGCAGCCTTGGTGTGACAACCTCCGTTTTCTTCTCCAGGTCTCGCTAGGGGTCTTTCCATTTATGACTCTCTTCA
CAGAAGAGTTTACAGTGTGCTGATTTCCGGCTGTTTCTGCGTAATTGGTGTCTGCTGTTATCGATGGCCTCCTTCCA
TTTCTTTTAGGCTTTGTTTATGTTGTTTTCCGGCTCCTTGAAGGAAAAGTTTCGATTATGGATGTTTGAACCTTTCTTT
30 TCTAAACAAGCATCTGAAGTTGCCGTTTTCCCTCTAAAGCAGGGATCCCGAGGCCCTGGCTGTGGAGTGGCACCGGTCT
GGGGCTGTAGGAACCCGGCGCACAGCGGGAGGCTAGGTGGGGTGTGGGAGCCAGCGTTCCCGCTGAGCCCGCCCC
TCTCAGATCAGCAGTGGCATGCGGTGCTCAGAGGCGCACACCCCTACTGAGAACTGTGCGTGAGAGGGGTCTAGATTCT
GTGCTCCTTATGGGAATCTAATGCCTGATGATCTGAGGTGGAACCGTTTGCTCCCAAACCATCCCTTCCCCACTGTG
TCCTGTGGAAAAATCGTCTTCCACGAAACCACTCCCTGGTACCACAATGGTTGGGGACCTGTGCTAAAGACCTGCTTCA
35 GCAGCCTCTCGTCAGTGTGATATATGGCTTTTCTGTGTTGAGTCCAGAATAATTACGGATTCTGTGATGCTTTCCGC
CGACCTCAGACCCATGGGCTATTTGTGGCGGTGTGCTGCTCCTGGGTGGGAAGGGTGACGGCCCCATGTACCTTCT
GTTACTGCCTTCCAGGTTGGTTCTCAGGGTTGAATCGTACTCGATGTGGTTTTAGCCCACGGCCCTGCCGCCAGCTCCTG
GGGGCTGGGGAACATGCTGAAGCACAGATCACCGTGCCTCTTTGATGCCTCACAAGCTCGAGGCCTCCTGTGTCGG
TGTTAGTGTGTGTCAGTGCCTGCTCACATCCTGTCTTGGGGACGCAGGGGCTTAGCAGGTCCCGTAGTAAATGACAAGC
40 GTCTGGGGGAGTCTGCAGAATAGGAGGTGGGGTGCCGTCTCTCTCCCGGTCTTCAGACTCTTCTCCTGCTGTGCT

GTGGCTGCACCTGCATCCCTGCAATCCCTCCAGCACTGGGCTGGAGAGGCCCGGGAGCTCGAGTGCCACTTGTGCCACGT
GACTGTGGATGGCAGTCGGTCACGGGGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTGGTACAGGGGTCTGATGTGTG
GTGACTGTGGATGGCGGTCTGATGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTG
ATGGCGGTCTGATGTGGTGAAGTGTGGATGGCGGTCTGATGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
5 GGGTCTGATGTGGTGAAGTGTGGATGGCAGTCGGGGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
TGGTGAAGTGTGGATGGCAGTCGGGGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
GTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
CGGTCTGATGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCA
CAGGGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCA
10 GGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
TCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
GATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
GACTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGT
GGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
15 TCGTGGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
TCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
GGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
GACTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
GGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
20 GTGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
GGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
GTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGTGTGGATGGCGGTCTGATGTGTGGTGAAGT
ACTTTGCGTCTCGGCCCCCGGCCCCGTTTCCCAAACAGAAGCTTCCAGGCGCTCTCTGGGCTTCATCCCGCATCG
GGCTTGGCCGAGGTCCACAGTCTGATCGGAAGAAACAAGTGCCAGCTCTGGCCGGGCGAGCCACATTTGTGGCTC
25 ATGCCCTCTCTGCGGCAG

Intron 7 (SEQ ID NO 11)

GTCTGGGCACTGCCCTGCAGGGTTGGGCAAGGACTCCAGCAGTGGGTCTCTCCCTGGGCAATCACTGGGCTCATGACCG
GACAGACTGTTGGCCCTGGGGGAGTGGGGGAATGAGCTGTGATGGGGCATGATGAGCTGTGTGCTTGGCGAAATC
30 TGAGCTGGGCCATGCCAGGCTGCGACAGCTGCTGCATTAGGCACCTGCTCACGTTGACTGCGCGGCTCTCTCCAGTT
CCGAGTGCCTTTGTTTCATGATTTGCTAAATGCTTCTCTGCCAGTTTGTATCTTGAGGCCAAAGGAAAGGTGTCCCT
CCTTTAGGAGGGCAGGCCATGTTTGGAGCGGTCTGCTGCCAGCTGGCCCTCAGTGTGGGTCTGAGGCCAAAGGAAAGG
TGTCCCTCTCTTAGGAGGACGGGCGGTGTTTGGAGCCAGCCCCGCTGAGCGGGCTCTCAGTGTGGGTCTGTCCACGT
GGCCCTGTGGCCCTTTGCAGATGTGGTCTGTCCAGTGGCCCTGTGGCTCTTTCAGATGCCTGTTAGCACTTGTCTGGC
35 TCTAGGGGACAGTCTGTCCACCGCATGAGGCTCAGAGACCTCTGGGCGAATTTCTTGGCTCCAGGGTGGGGTGGAG
GTGGCTGGGCTGTGGGACCCAGACCTGTGCCCCGCGAGCTGGGCGCAACTCTGGATCACATATGCCATCCGGGCCA
CGGTGGGCTGTGTGGGTGTGAGCCAGCTGGACCCACAGGTGGCCCGAGGAGACGTTCTGTGTACACACTCTGCCTAA
GCCCATGTGTCTGACAGACTCGGCCCGGCCAGCCACAGTGGCCCTGCATTCCAGCCAGCCCGCACTTCATCACA
AACACTGACCCCAAAGGGACGGAGGGTCTTGGCCAGTGGTCTGCTGTCTCAGCACCACCGGCTCACTCCCATGTG
40 TCTCCCGTCTGCTTTCGAG

5

10

15

20

25

30

35

40

CGTTAGGGTCCTTGGGGAGATGGGGCTGGTGAGCCTGAGGCCCCACATCTCCAGCAGGCCCTCGACAGGTGGCCTGGA
CTGGGCGCCTCTTCAGCCCATTCGCCATCCCACTTGCAATGGGGTCTACACCAAGGACGCACACACCTAAATATCGTGCC
AACCTAATGTGGTTCAACTCAGCTGGCTTTTATTGACAGCAGTTACTTTTTTTTTTTAATACTTTAAGTTCTAGGGTAC
ATGTGCACGACGTGCAGGTTAGTTACATATGTATACATGTGCCATGTTGGTGTGTGTCACCCATTAACTCATCATTTACA
5 TTAGGTATATCTCCTAATGCTATCCCTCCCCACTCCCCCATCCCATGACAGGCCCTGGTGTGTGATGTTCCCCACCCTG
TGCCAAGTGTTCTCATTGTTCAAGTCCCACCTGTGAGTGAGAACATGTGGTGTGTTGGTTTTCTTTCTCTGCAATAGTTT
GCTCAGAGTGATGGTTTCCAGCTTCGTCCATGTCCCTACAAAGGACATGAACTCATCTTTTTTATGACTGCATAGTATT
CCGTGGTGATATGTGCCACATTTTCTTAATCCAGTCTATCATCGATGGACATTTGGGTGGTTGCAAGTCTTTGCTACT
GTGAATAGTGCCGCAATAAACATACGTGTGCATGTGTTTATAGCAGCATGATTATAATCCTTTGGGTATATACCCAG
10 TAATGGGATGGCTGGGTCAAATGGTATTCTAGTTCTAGATCCTTGAGGAATCACCACACTGTCTTCCACAATGGTTGAA
CTAGTTTACACTCCCACCAACAGTGTAAGTGTTCTGGTGCTGGAGAGGATGTGGACAGCAGTTATTTTTTTATGAAAA
TAGTATCACTGAACAAGCAGACAGTTAGTGAAGGATGCGTCAGGAAGCCTGCAGGCCACACAGCCATTTCTCTCGAAGAC
TCCGGGTTTTCTCTGTGCATCTTTTGAAACTTAGCTCCAATTATAGCATGTACAGTGGATCAAGGTTCTTCTTCATTAA
GGTTCAAGTTCTAGATTGAATAAGTTTATGTAAACAGAAACAAAAATTTCTGTACACACAACCTGTCTGGGATTTGGA
15 GGAAAGTGTCCTCGAGCTGGCGGCACACTGGTCAGCCCTCTGGGACAGGATACCTCTGGCCCATGGTCATGGGCGCTGG
GCTTGGGCTGAGGGTCACACAGTGACCATGCCAGCTTCCTGTGGATAGGATCTGGGTCTCGGATCATGCTGAGGACC
ACAGCTGCCATGCTGGTAAAGGGCACCAGTGGCTCAGAGGGGGCGAGGTTCCAGCCCCAGCTTCTTACCGTCTTCAG
TTATTTTTCCCTAAGAGTCTGAGAAGTGGGGCCGCGCTGATGGCTTCGTTCTGTCTTCTCAGCTGGCACAGAATTGCACAA
GCTGATGGTAAACACTGAGTACTTATAATGAATGAGGAATTGCTGTAGCAGTTAACTGTAGAGAGCTCGTCTGTGGAAA
20 GAAATTTAAGTTTTTCAATTAACCGCTTGGAGAATGTTACTTTATTATGGCTGTGTAATGTTTGACATTAGTCC
TCGTAGACAGATACTACGTAAAAAGGTAAAGTTAACCTTGCTGTGATTTTCCCTTATTTAG

Intron 10 (SEQ ID NO 14)

GTGAGGCCCGTGCCGTGTGTCTGTGGGACCTCCACAGCCTGTGGGCTTTGCAGTTGAGCCCCCGTGTCTGCCCCCTGG
25 CACCGCAGCGTTGTCTCTGCCAAGTCTCTCTCTGCGGTGTGGATCCGCAAGAGCAGAGGCGCTTGGCCGTGCACC
CAGGCTCGGGGCGCAGGGGCACCTTCGGGAGGGAGTGGGTACCGTGAGGCCCTGGTCTGACAGAGCGCACCCAGGTT
ACACAGTGGTGAGTGACGGCGTGACCTGGCTCCTGCTGCTCTTGGAAAGTCAAGAGTGGCGGCTCCTGGGGCCCCAG
TGAGACCCCGAGAGCTGTGCACAGGGCTGCAGGGCCGAGGCGGCAGCCTCCTCCCGAGGGTGACCTGAGCCTGCGGA
GAGCAGGAGCTGCTGAGTGAGCTGGCCACAGCGTTCTGCTGCGGTACGTTCTGCGTGGGGTTGTTTGGGATCGGTGGG
30 AGAATTTGGATTTGCTGAGTGCTGCTGTCTTGAACCACGGAGATGGCTAGGAGTGGGTTTCAGAGTTGATTTTGTGAAT
CAAATAAAATCAGGCACAGGGGACCTGGCCTCAGCACAGGGGATTGTCCAATGTGGTCCCCCTCAAGGGCGCCCCACAG
AGCCGCTGGGCTTGTTTTAAAGTGCGATTGACGAGGGACGAGAAACCTTGAAAGCTGTAAAGGAACCTCAGAAAATG
TGGCCGCCAGGGGTGGTTTCAGGTGCTTTGCTGGGCTGTGTTTGTGAAAACCATTTTGACCCGCCCTCCAAGTCCACCC
TCCAGGTCCACCTCCAGGGCCGCCCTGGGCTGGGGTATGCTGGCGTTCTTGTGCCGAGCCCGAGCACAGCAGGC
35 TGTGCACATTTAAATCCACTAAGATTCACTCGGGGGAGCCAGGTCCCAAGCAACTGAGGGCTCAGGAGTCTGAGGCT
GCTGAGGGGACAGAGCAGACGGGGAACGCTGCTTCTGTGTGGCAAGTTCCTGAGGGTGTGGCCAGGGAGGTGGCTCAGA
GTGTATGTTGGGTTCCACCGGGGGCAGAACTCTGTCTCTGATGAGTCGGCAGCCATGTAACAGGAAGGGGTGGCCACAG
GGAGCTGGGAATGACACAGGGGAGCTGCGCAGCTGGCCGAGGTCCAGGGCCAGGCCACAGGAAGGGCAGGGGGACGCCC
GGGGCCACAGCAGAGGCCCGCAGGAAGGGAAGGGATGCCAGGCCAGAGCAGAGGCTACCGGGCACAGGGGGCTCCCTG
40 AGCTGGGTGAGCGAGGCTCATGACTCGGCGAGGGAACCTCCTTGACGTGAAGCTGACGACTGGTGTGCCAGCTCACAG

5 CCCAGCCAGGTCCCGCGCTGAGCAGGAACTCAGAACCCTCCCTTTGTCTAAAGCACAGCAGATGCCTTCAGGGCATCT
AGGAGAAAACAGGCAAAGTCGTTGAGAAACGTCCTTAAAGAAGGTGGGATGGTGGCAATTTCTGTCCAGATTTTAGTCT
GCCCCGACCACAGATGAGTCTATAACGGGATTGTGGTGTGCCATGGGGACACATGAGATGGACCATCACAGAGGCCAC
TGGGGCTGCACCTCCCATCTGAGTCTGGCTGTCCCGGGTCCAGGCCAGGTTCTTGCTATGCTACCTACCTGTCTGTCCC
GGGAGACAGGGAAGCAGCCCGAAGCTGGAGCAGGGCTGGGTCCAGGCTCCTCAGAGCTCCTGCCAGGCCAGCACCTT
GCTCCAAATCACCCTTCTCTGGGGTTTTCCAAAGCATTAACAAGGGTGTGAGTTACCTCCTGGGTGACGGCCCCGCA
TCTTGGGGCTGACATTGCCCCCTCTGCCTTAG

Intron 11 (SEQ ID NO 15)

10 GTGAGCGCACCTGGCCGGAAGTGGAGCCTGTGCCCGGTGGGGCAGGTGCTGCTGCAGGGCCGTTGCGTCCACCTCTGCT
TCCGTGTGGGGCAGGCGACTGCCAATCCCAAAGGTGAGAGGCCACAGGGTGCCCTCGTCCCATCTGGGGCTGAGCAGA
AATGCATCTTTCTGTGGAGTGAGGGTGCTCACAACGGGAGCAGTTTTCTGTGCTATTTTGGTAAAGGAAATGGTGAC
CAGACCTGGGTGCACTGAGGTGTCTTCAGAAAGCAGTCTGGATCCGAACCCAAGACGCCCGGGCCCTGCTGGGCGTGAGT
CTCTCAAACCGAACACAGGGGCCCTGCTGGGCATGAGTCCCTCTGAACCCGAGACCCTGGGGCCCTGCTGGGCGTGAGT
15 CTCTCCGAACCCAGAGACTTCAGGGCCCTTTTGGGCGTGAGTCTCTCCGCTGTGAGCCCCACACTCCAAGGCTCATCCAC
AGTCTACAGGATGCCATGAGTTCATGATCAGTGTGACCCATCAGGGGACAGGGCCATGGTGTGGGGGGGTCTCTACAA
AATTCTGGGGTCTTGTTCCTCCAGAGCCGAGAGCTCAAGGCCCGCTCTCAGGCTCAGACACAAATGAATTGAAGATGGA
CACAGATGCAGAAATCTGTGCTGTTTCTTTATGAATAAAAAGTATCAACATTCAGGCAGGGCAAGGTGGCTCACACCT
ATAATCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACTTGAGGCCAGGAGTTGAGGCCAACCTAACCAACATAGTG
20 AAATTCATTTCTACTTAAAAAATACAAAAATTAGCCTGGCCTGGTGGCACACGCCTGTAGTCCCCGCTATGCGGGAGGC
TGAGGCAGGAGAATCATTGAACCCAGGAGGCAGAGGTGTCAGTGAGCCGAGATCACACCACTGCACTCCAGCCTGGGCA
ACAGAGTGAGACTTCATCTTAAAAAAGAGTATCAGCATTCCAAAACCATAGTGAGCAGGTGTTTTTTTATTC
TGCTCTTCGATAATATTTACTGGTGTGTGTAGAGGCCGGAACCTGGGGGTGCCTTCCTCTGAAAGGCACACCTTCATGG
GAAGAGAAATAAGTGGTGAATGGTTGTAAACCAGAGGTTTAACTGGGGTCTGTGCTTCTGAGTTAACAGTCCAGATC
25 TGGACTTTGCCTCTTTCCAGATGCTCCCTGGGGTTTGCTTCATGGGGAGCAGCAGGTGTGGACACCCTCGTGATGGGG
GAGCAGCAGGTGCAGACGCCCTCATGATGGGGAGTGGCAGGTGCAGACACCCTTGTGATGGTGGCCAGCATGCTCCCTG
TTGCAGCTCCCTCCCCACAAGGATGCCGGTCTCCTGTGTCTCCCAAGTCCCTGCTTCCCTCTCACAGCCTTACCTGGTC
CTGGCCTCCACTGGCTTTGTCTGATGATTTCCACATTTCTGGGCTCCAGCACCTCTTCGCCTCTCCAGGCACCTCT
GCAGTGCTGGCCATACCACTGAGCTGTGAAGTGTCCACTGCTTATTTTGCTCCCATGAAATGTATTTTATAGGACAGGC
30 ACCCCTGGTTCCAGCCTCTGGCACAGCATCAGTGAATGTTATTGAAGGACAAAGGACAGACAAACAAATCAGGAAAATGG
GTTCTCTCTAAACACATTGCAAAGCCACAGAGGCTAGTGCAGGATGGGTGGGCATCAGGTATCAGATGTGGGTCCAATG
CCAGAATATTCTGTGCTCCCAAAGGCCACTTGGTCAGAGTGTGTGCTGTCAGAGGTGGCTCTAAAAGCTCAGCAGTGGAG
GCAGTGGTTCCGCATACTCAGGGTGAAGTACATCCTCTGTGTCTGAAGTATACAGCAGAGGCTTGAAGGGCATCTGGGA
GAAGAAAACAGGCAAATGATTAAAGAAAAGTAAAAAGGAAAAGTGGTAAGATGGGAATTTCTTGTCCAGATTTTAGTC
35 TCCCAAACCAAGCTCAGATGGTAGAATGTGGTCAGAACTGATGGACAGAAACATAGAACAAACGGAAGCCCTATCTCT
CAGAAACGTGTGTTAATGTGGTATGTGGCACAGCTGATGGAAGAGAGTGTGTGTGTAATTTTTTTTCTGAGAAAAT
GACTGGAAGCAAATAAGTTGTGCTTTTACAGCATATACCAGAGCAGATTCTAGGTAGAAGAGGAGACATGCAAAACAC
ACCAAGCAACAGAAATAAAACAAAAGACTCAAAGGGAAGGGAGGTGAACGTTCCCTGGTTTGGTGTGGGGAAGGACACAC
AGGGAGGCGGATGAAACCAAGTGAGGCAACGGGCATTGCTTTCACTGCAGAGAACTCAGCTTGCTGAGCCACAGTGAAA
40 ATGGCCATTCCCTGGAGCGTTTGTGACGTGATTTATTTAAGGCGCCCTGTGAGGTCCTGCACATTCATCTCTCACTTT

5 GTTCTCCTAACACCTGAGAGGTAGAGGAGGAAAGGCTCCAGGGGAGCAGCCGCCCTTGGTCACCCAGCTGGCAAAGGGC
ATGCATGATTGCAGCCTGGCCTCCTGCTCCGGGGCCCTTGCTCTGCCCGAGGACCCACACAAGTCAGACCCATAGGCTC
AGGGTGAGCCGGAGCCCAAGGTCGTGTTGGGGATGGCTGTGAAAGAAGAAATGGACGTCCTGATGCACACTTGGGAAGGTC
CTACCAGCAGCGTCAAAGAAATGCATGTGAAACTGACAGCGAGACCCATCCCTCAAAGAAACGCACGTGAAACTGATGGC
10 GAGACCTGTCCCATCCCTCATGCTGGCTCCTTTCTGGGCTTGCCAAGAGCCAGCATCAGGTTGAGGCAAGCTGGAAAG
ACTTTCTGGAAAGCAGCTTGTTCATGGAAGTCCTCACAATGTCTGTGCTTCCAGTAATTCACCTTCTGAAGTGA
CCAGACATTATCACGGGTCTTATTTACCATTTCAGTGTTCCAGGCAGGGGACTTGCCACAGCAAGTCACGAACCTGCC
CAAATACAGGGCTAAGGAGATATTATGCATCAGAAACTTGCTCTGCCATTAAACATTTTCAAAGAATTTTGAAGAAT
GTTTAATGGCAGAAACGTTTATTTCATGTAGCAGTGTTCAAAGCTGGATGTAAGAAGACACACCCAGGAGCCTGCCG
15 TGAATGTATGTGTGTTTCATCTTTGACATGGACATACATGGGCAGTGAGTGGTGGTGAGGCCCTGGAGGACATCGGTGG
GATGCCTCCATCCTGCCCTCTGGAGACCATGTGTGCCAGTGCACTCACTGGAGCCCTGTTAGCTGGTGCCACCTG
GCTCTTCCATCCCTGAGATTCAAACACAGTGAGATTCCCCACGCCCAACTCAGTGTTCTCCCAAAAAACCTGAGTCAC
ACCTGTGTTCACTCGAGGGACGCCCGGGAGCCAGGGCTCCACAGTTTATTATGTGTTTTGGCTGAGTTATGTGCAGATC
TCATCAGGGCAGATGATGAGTGACAAACACGGCCGTGCGAGGTTTGGATACACTCAACATCACTAGCCAGGTCCTGGTG
20 GAGTTTGGTCATGCAGAGTCTGGATGGCATGTAGCATTTGGAGTCCATGGAGTGAGCACCAGCCCCCTCGGGCTGCAGC
GCATGCCCCAGGCAGGACAAGGAAGCGGGAGGAAGGCAGGAGGCTCTTTGGAGCAAGCTTTGAGGAGGGGGCTGGGTGT
GGGCGAGGCACCTGTGTCTGACATTCCCCCTGTGTCTCAG

Intron 12 (SEQ ID NO 16)

20 GTGAGCAGGCTGATGGTCAGCACAGAGTTCAGAGTTCAGGAGGTGTGTGCGCAAGTATGTGTGTGTGTGTGCGCGCGT
GCCTGCAAGGCTGATGGTGACTGGCTGCACGTAAGAGTGACATGTACGCATATACACGTGAGCACATACATGTGTGCAT
GTGTGTACATGAAGGCATGGCAGTGTGTGCACAGGTGTGCAAGGGCACAAGTGTGTGCACATGCGAATGCACACCTGACA
TGATGTGTGTTCTGTGCACAGTCGTGTGGGCATTACGTGAGGTGCATGCGTGTGGGTGTGCAGTGTGAGTAGCATGTGT
GCACATAACATGTATTGAGGGGTCTCGTGTTCACCCCGTAGTCTCTCAGCACCAGTGCCACTCCTTACAGGATGAGAC
25 GGGGTCCCAGGCCTTGGTGGGCTGAGGCTCTGAAGTGCAGCCCTGAGGGCATTGTCCCATCTGGGCATCCGCGTCCACT
CCCTCTCCTGTGGGCTTCTGTGTCCACTCCCCCTCTCCTGTGGGCATTACATCCACTCCACTCCCTCTCTCTGTGGGC
ATCCGCGTCCACTCCCCCTCTCTGTGGGCATCTGCGTCCACTCCCCCTCTCTGTGGGCATTGCGTCCACTCCCTCTCCT
GGTTCCTTCTGTCTTGGCCGAGCCTCGGGGCAGGCAGATGACACAGAGTCTTGACTCGCCAGGGTGGTTTCGAGCTG
CCGGGTGAGGGCCAGGCCGATTCTCACTGGGAAGAGGGATAGTTTCTTGTCAAATGTTCTTTCTTGTTCATCTGA
30 ATGGATGATAAAGCAAAAAGTAAAACTTAAATCCAGAGAGGTTTCTACCGTTTCTCACTCTTTCTTGGCGACTCTAG

Intron 13 (SEQ ID NO 17)

35 GTGAGCCGCCACCAAGGGGTGAGGCCAGCCTCCAGGGACCCTCCGCGCTCTGCTCACCTCTGACCCGGGGCTTCACTT
TGGAACCTCTGGGTTTTAGGGGCAAGGAATGTCTTACGTTTTTCACTGGTGCTGCTGCCTGTGCACAGTTCTGTTTCGCGTG
GCTCTGTGCAAGCACCTGTTCTCCATCTCTGGGTAGTGGTAGGAGCCGGTGTGGCCCCAGGTGTCCCACTGTGCCTGT
GCACTGGCCGTGGGACGTATGGAGGCCATCCAGGGCAGCAGGGGCATGGGGTAAAGAGATGTTTATGGGGAGTCTTAG
CAGAGGAGGCTGGGAAGGTGTCTGAACAGTAGATGGGAGATCAGATGCCCGAGGATTTGGGGTCTCAGCAAAAGAGGGCC
GAGGTGGGTGCAGGTGAGGGTCTGCGCCCCACCCCGGAAGGTGCAGCAGAGCTGTGGCTCCCCACACAGCCCGGCCA
GCACCTGTGCTCTGGGCATGGCTGTGCTCCTGGAACGTTCCCTGTCTGGCTGGTCAGGGGGTCCCTGCCAAGAATCG
40 ACAACTTTATCACAGAGGGAAGGGCCAATCTGTGAGGCCACAGGGCCAGCTTCTGCCTGGAGTCAGGGCAGGTGGTGGC

ACAAGCCTCGGGGCTGTACCAAAGGGCAGTCGGGCACCACAGGCCCGGGCCTCCACCTCAACAGGCCTCCCGAGCCACTG
GGAGCTGAATGCCAGGAGGCCGAAGCCCTCGCCCATGAGGGCTGAGAAGGAGTGTGAGCATTGTGTACCCAGGGCCG
AGGCTGCGCGAATTACCGTGACACTTGATGTGAAATGAGGTCGTCTATCGTGAAACCCAGCAAGGGCTCACGGGA
GAGTTTTCCATTACAAGGTCGTACCATGAAAATGGTTTTTAACCCGAGTGTGCGCCTTCATGCTCTGGCAGGGAGGGC
5 AGAGCCACAGCTGCATGTTACCGCCTTTGCACCAGCTCCAGAGGCTTGGGACCAGGCTGTCTCAGTTCAGGGTGCCTCC
GGCTCAGACCGCCCTCCTCTCTGCTTCTCTCTCTGCTCAAATCTTCCCTCGTTTGCATCTCCCTGACGCGTGCTGGG
CCCTCGTGCAAGCTGCTTGACTCCTTTCCGGAACCCCTTGGGGTGTGCTGGATACAGGTGCCACTGAGGACTGGAGGTGT
CTGACACTGTGGTTGACCCAGGGTCCAGCTGGCGTGCTGGGGCCTCCTTGGGCCATGATGAGGTGAGGAGAGTTTTCC
CAGGTGAAAACCTCCTGGGAACTCCAGGGCCATGTGACCTGCCACCTGCTCCTCCCATATTAGCTCAGTCTTGTCTCTC
10 ATTTCCACCAGGGTCTCTAGCTCCGAGGAGCTCCCGTAGAGGGCCTGGGCTCAGGCGAGGCGGCTGAGTTTCCCCAC
CCATGTGGGGACCTTGGGTAGTCGCTTGATTGGGTAGCCCTGAGGAGGCCGAGATGCGATGGGCCACGGGCCGCTTCCA
AACACAGAGTCAGGCACGTGGAAGGCCAGGAATCCCTTCCCTCGAGGCAGGAGTGGGAGAACGGAGAGCTGGGCCCCG
ATTTACGGCAGCCAGGCTGCAGTGGGCGAGGCTGTGGTGGTCCACGTGGCGCTGGGGCGGGGTCTGATTCAAATCCGC
TGGGGCTCGGCCTTCTTGGCCCCTGTGCGCGCCTCCACACGGGCTTGGGGTGGACGCCCGACCTCTAGCAGGTGGC
15 TATTTCTCCCTTTGGAAGAGAGCCCTCACCCTAGCTAGGTGTTTCCCTCCTGGGTGAGGAGCGTGGCCGTGTGGCAACC
CCGGGACCTTAGGCTTATTTATTTGTTTAAAAACATTCTGGGCTGGCTTCGGTTGTTGCTAAATGGGGAAAAGACATCC
CACCTCAGCAGAGTTACTGAGAGGCTGAAACCGGGTGTGGCTTGACTGGTGTGATCTCAGGTCAATCCAGAAGTGGCT
CAGGAAGTCAGTGAGACCAGGTACATGGGGGCTCAGGCAGTGGGTGAGTGAGGTACACGGGGGCTCAGGCAGTGGGT
GAGGCCAGGTACATGGGGGCTCAGGCACTGGGTGAGATGAGGTACACGGGGGCTCAGGCAGAGGCTCAGACCAGGTAC
20 ACGGGGGCTCTGATCACACGCACATATGAGCACATGTGCACATGTGCTGTTTCATGGTAGCCAGTCTGTGCACACCTGC
CCCAAAGTCCAGGAAGCTGAGAGGCCAAAGATGGAGGCTGACAGGGCTGGCGGGTGGCTCACACCTGTAGTCCAGCA
CTTTGGGAGGCCGAGCGAGAGGATCCCTTGAAGCCAGGAGTTAAGACCAGCCTGAGCAACATAGTAGAACCCCATCTC
TATGAAAAATAAAAAAAAATTAGCTGAACATGGTGGTGTGCGCCTGTAGTTCCAATACTTGGGAGGCTGAAGTGGGAG
GATCACTTGAGCCCAGGAGGTGGAAGCTGCAGTGAGCTGAGATTGCACCACTGTACTGCAGCCTGGGTGACAGAGTGAGA
25 GCCCATCTCAACAACAACAAGAAGACTGACAAATGCAGTTTCTTGGAAAGAAACATTTAGTAGGAACCTTAACCTACACA
CAGAAGCCAAGTCGGTGTCTCGGTGTGAGTGTGATGATGGGTCTCACACCATCACCCAGACCCAGGGTTTATG
CACCACAGGGGCGGTGGCTCAGAAGGATGCGCAGGACGTTGATATACGATGACATCAAGTTGTCTGACGAAGGGCAG
GATTCATGATAAGTACCTGTGGTACACAAGGAACAATGGATAAACTGGAAACCTTAGAGGCCTTCCCGGAACAGGGGCT
AATCAGAAGCCAGCATGGGGGCTGGCATCCAGGATGGAGCTGCTTCAGCCTCCACATGCGTGTTCATACAGATGGTGCA
30 CAGAAACGCAGGTACCTGTGCACACACAGACACGAGCTACTCGCACACAAGCACACACAGACATGCATGCATGC
ATCCGTGTGTGCACCTGTGCCCATGAGGAAACCATGTCATGTGCATTATGCACGCACACAGGCACCGGTGGGCCCAT
GCCACACCCACGAGCACCGTCTGATTAGGAGGCCCTTCTCTGACGCTGTCCGCCATCCTCTCAG

Intron 14 (SEQ ID NO 18)

GTATGTGCAGGTGCCTGGCCTCAGTGGCAGCAGTGCCCTGCTGCTGTTAGTGTGTGAGGAGACTGAGTGAATCTGGG
35 CTTAGGAAGTTCTTACCCCTTTTCGCATCAGGAAGTGGTTTAACCCCAACCACTGTGAGGCTCGTCTGCCCCCCTCTCGT
GGGGTGAAGCAGACACCTGATGGAAGGACAGGAGCTGTCTGGGAGCTGCCATCCTTCCACCTTGCTCTGCTGGGGAA
GCGCTGGGGGCTGGTCTCTCTGTTTGGCCCATGGTGGGATTGGGGGCTGGCCTCTCTGTTTGGCCTGTGGTGG
GATTGGGCTGTCTCCGTCCATGGCACTTAGGGCCCTTGTGCAAACCCAGGCCAAGGGCTTAGGAGGAGGCCAGGCCAG
40 GCTACCCACCCCTCTCAGGAGCAGAGGCCGCTATCACCACGACAGAGCCCGCGCGCTCTGCTTCCAGTCACCG

TCCTCTGCCCCCTGGACACTTTGTCCAGCATCAGGGAGGTTTCTGATCCGTCTGAAATTCAAGCCATGTGAACTGCGGT
CCTGAGCTTAACAGCTTCTACTTTCTGTTCTTCTGTGTGTGGAAATTTACCTGGAGAAGCCGAAGAAAACATTCTCG
TCGTGACTCCTGCGGTGCTTGGGTGCGGACAGCCAGAGATGGAGCCACCCCGCAGACCGTCCGGTGTGGGCAGCTTTCCG
GTGTCTCCTGGGAGGGGAGCTGGGCTGGGCTGTGACTCCTCAGCCTCTGTTTTCCCCCAG

5

Intron 15 (WEQ ID NO 19)

GCAAGTGTGGGTGGAGGCCAGTGCAGGGCCCCACCTGCCCAGGGGTCTCCTTGAACGCCCTGTGTGGGCGAGCAGCCTC
AGATGCTGCTGAAGTGCAGACGCCCCCGGGCTGACCTGGGGGCTGGAGCCACGCTGGCAGCCCTATGTGATTAAACG
CTGGTGTCCCCAGGCCACGGAGCCTGGCAGGGTCCCCAAGTTCTTGAACCCCTGCTTCCCATCTCAGGGGCGATGGCTCC
CCAGCCTTGGGAGCCTTCTGACCCCTGACCTGTGTCTCTCACAGCCTCTTCCCTGGCTGCTGCCCTGAGCTCCTGGGGT
CCTGAGCAAGTTCTCTCCCGCCCCGCGCTCCAGCGTCACTGGGCTGCCTGTCTGCTCGCCCGGTGGAGGGGTGTCTG
TCCCTTCACTGAGGTTCCACAGCCAGGGCCACGAGGTGCAGGCCCTGCCTGCCCGCCACCCACACGCTCTAGGAGGG
TTGGAGGATGCCACCTCTGGCCTCTTCTGGAACGGAGTCTGATTTTGGCCCCGAG

10

15

3'-untranskribierter Bereich (SEQ ID NO 20)

ATCTCATGTTTGAATCCTAATGTGCACTGCATAGACACCACTGTATGCAATTACAGAAGCCTGTGAGTGAACGGGGTGGT
GGTCAGTGCAGGGCCCATGGCCTGGCTGTGCATTACGGAAGTCTATGAGTGAATGGGGTTGTGGTCAGTGCAGGGCCCATG
GCCTGGCTGGGCTGGGAGGTTTCTGATGCTGTGAGGCAGGAGGGGAAGGAGGGTAGGGGATAGACAGTGGGAGCCCCCA
CCCTGGAAGACATAACAGTAAGTCCAGGCCCCGAAGGGCAGCAGGGATGCTGGGGGCCAGCTTGGGCGGCGGGGATGATG
GAGGGCCTGGCCAGGTTGGCAGGGATGATGGGGGCCACAGCTGGGGTGGCAGGGGTGATGGGGGGGGTGGTCTGGGTGG
CGGGGAAGATGGGAAGCCTGGCTGGGCCCCCTCTCCCTGCCTCCACCTGCAGCCGTGGATCCGATGTGCTTCCCT
GGTGACATCCTCTGGGCCATCAGCTTTCATGGAGGTGGGGGGCAGGGGCATGACACCATCCTGTATAAAATCCAGGATT
CCTCCTCTGAACGCCCAACTCAGGTTGAAAGTCAATTCCGCCTCTGGCCATTCTCTTAAGAGTAGACCAGGATTCTG
ATCTCTGAAGGGTGGGTAGGGTGGGGCAGTGGAGGGTGTGGACACAGGAGGCTTCAGGGTGGGGCTGGTGATGCTCTCTC
ATCCTCTTATCATCTCCAGTCTCATCTCTCATCCTCTTATCATCTCCAGTCTCATCTGCTTCTCTTATCTCCAGT
CTCATCTGTATCCTCTTACCATCTCCAGTCTCATCTCTTATCCTCTTATCTCCTAGTCTCATCCAGACTTACCTCCA
GGGCGGGTGCAGGCTCGCAGTGGAGCTGGACATACGTCCTTCCCTCAGGCAGAAGGAACTGGAAGGATTGCAGAGAACAG
GAGGGGCGGCTCAGAGGGACGCAGTCTTGGGGTGAAGAAACAGCCCCCTCTCAGAAGTTGGCTTGGGCCACACGAAACCG
AGGGCCCTGCGTGAGTGGCTCCAGAGCCTTCCAGCAGGTCCCTGGTGGGGCCTTATGGTATGGCCGGGTCCTACTGAGTG
CACCTTGGACAGGGCTTCTGTTTGTAGTGCAGCCCGACGTGCCTGGTGTGGGGTGGGGCTTATGGCCACTGGATATG
GCGTCATTATTGCTGTGCTTCCAGAGAATGTCTGAGTGACCGAGCCTAATGTGTATGGTGGGCCAAGTCCACAGACTG
TGTCGTAAATGCACTCTGGTGCCTGGAGCCCCGTATAGGAGCTGTGAGGAAGGAGGGGCTCTTGGCAGCCGGCCTGGGG
GCGCCTTTGCCCTGCAAACTGGAAGGGAGCGGCCCGGGCGCGTGGGCGGACGACCTCAAGTGAGAGGTTGGACAGAAC
AGGGCGGGGACTTCCCAGGAGCAGAGGCCGCTGCTCAGGCACACCTGGGTTTGAATCACAGACCAACAGGTCAGGCCATT
GTTCAGCTATCCATCTTCTACAAAGCTCCAGATTCTGTTTCTCCGGGTGTTTTTTGTTGAAATTTTACTCAGGATTACT
TATATTTTTTGCTAAAGTATTAGACCTTAAAAAAGGTATTTGCTTTGATATGGCTTAACCTACTAAGCACCTACTTTAT
TTGTCTGTTTTTATTATTATTATTATTATTATTAGAGATGGTGTCTACTCTGTCAACCCAGGTTGTTAGTGAGTGGCAC
AGTCATGGCTCGCTGTAGCCGCAACCCCCAGGCTCAAGTGATCCTCCGGCCTCAGCTTCCAGAGTGTGGGATTACAG
GTGTGAGCCACTGCCCTTGCCTGGCACTTTAAAAACCACTATGTAAGGTCAAGTCCAGTGGCTTCCACACCTGTCTATCC
CAGTACTTTGGGAAGCCGAGGCAGAGGATTGTCTGAGGCCAGGAGTTTGTAGACCAGCATGGGTAACATAGGGAGACCCC

20

25

30

35

40

ATCTCTACAAAAATGCAAAAAGTTATCCGGGCGTGGGGTCCAGCATCTGTAGTCCCAGCTGCTCGGGAGGCTGAGTGGG
AGGATCGCTTGAGCCCGGGAGGTCAATGGCTGCAGTGAGCTGTGATTGTACCATCGCACTCCAGCCTGGGCAACAGAGTGA
GACCCCTGTCTCAAAAAAAAAAAAAAAAAAAGAGGAGAAGGAGAAGAGAAGAGAAGGAAGGAAGGAAGGAAGGAAGGAAG
5 GAAGAAGGAAGAAAGAGGAGAAGGAGGCTGTAGGTGCTAGGTAGACTGTCAAATCTCAGAGCAAAAATGAAAAATAACA
AAGTTTTAAAGGGAAGAAAAACCCAGCTCTTTGGACTTCCTTAGGCCTGAACCTTCATCTCAAGCAGCTTCCTTCCACA
GACAAGCGTGTATGGAGCGAGTGAGTTCAAAGCAGAAAGGGAGGAGAAGCAGGCAAGGGTGGAGGCTGTGGGTGACACCA
GCCAGGACCCCTGAAAGGGAGTGGTTGTTTTCTGCCTCAGCCCCAGCTCCTGCCGGTCTGACCTGCTGTAACCGTC
GATGTTGGTGCCAGGTGCCACCTGGGAAGGATGCTGTGCAGGGGGCTTGCCAACTTTGGTGGGTTTCAGAAGCCCCAG
10 GCACTTGTGGCAGGCACAATTACAGCCCTCCCCAAAGATGCCACGTCCTTCTCCTGGAACCTGTGAATGTGTACCCG
CAAGGCAGAGGCTGGTGAAGGCTGCAGGTGGAATCACGGCTGCCAGTCAGCCGATCTTAAGGTCATCTGGATTATCTGG
TGGGCCTGATATGGCCACAAGGGTCCCTAGAAGTGAGAGAGGGAGGCAGGGGAGAGTCAGAGAGGGGACGTGAGAAGGAC
CACTGGCCACTGCTGGCTTTGAGATGGAGGAGGGGGTCCCCAGCCAAGGAATGGGGGCAGCCGCTCCATGCTGGAAGGC
AAGCAATCCTCCCCGGTCTGAGGGCACACGGCCCTGCCACGCCTCGATTTCAGGCCAGTGGGACCTGTTTCAGCTTTT
15 CGGCCTCCAGAGCTGTAAGATGATGCGTTTGTGTTAGCCACTAAGCTGCAGTGATTGTCACAGCAGCAAATGGAATAG
CAGTACAGGGAAATGAATACAGGGACAGTTCTCAGAGTGACTCTCAGCCACCCCTGGG

Die Charakterisierung der Exons zeigte interessanterweise, daß die in unserer Patentanmeldung PCT/EP/98/03469 beschriebenen, funktionell wichtigen hTC-Protein-Domänen auf separaten Exons angeordnet sind. Das Telomerase-charakteristische T-Motiv befindet sich auf Exon 3. Die für die katalytische Funktion der Telomerase wichtigen RT (Reverse-Transkriptase)-Motive 1-7 liegen auf folgenden Exons: RT Motiv 1 und 2 auf Exon 4, RT Motiv 4 auf Exon 9, RT Motiv 5 auf Exon 10, RT Motiv 6 und 7 auf Exon 11. RT Motiv 3 liegt verteilt auf Exon 5 und 6 vor (s. Fig. 8).

10

Die Aufklärung der Exon-Intron-Struktur des hTC-Gens zeigt auch, daß die in unserer Patentanmeldung PCT/EP/98/03469 beschriebenen vier Deletions- bzw. Insertions-Varianten der hTC-cDNA ebenso wie drei weitere, in der Literatur (Kilian et al., 1997) beschriebene hTC-Insertions-Varianten höchstwahrscheinlich alternative Splice-Produkte darstellen. Wie in Fig. 8 gezeigt, lassen sich die Splice Varianten in zwei Gruppen einteilen: Deletionsvarianten und Insertionsvarianten.

15

Den hTC-Varianten der Deletionsgruppe fehlen spezifische Sequenzabschnitte. Die 36 bp in frame Deletion in Variante DEL1 resultiert höchstwahrscheinlich aus der Benutzung einer alternativen 3'-Splice Akzeptorsequenz in Exon 6, wodurch ein Teil des RT Motivs 3 verlorenght. In Variante DEL2 werden die normalen 5'-Splice Donor- und 3'-Splice-Akzeptor Sequenzen von Intron 6, 7 und 8 nicht benutzt. Stattdessen wird Exon 6 direkt an Exon 9 fusioniert, wodurch eine Verschiebung des offenen Leserahmens entsteht und in Exon 10 ein Stopcodon auftritt. Variante Del3 stellt eine Kombination aus Variante 1 und 2 dar.

20

25

Die Gruppe der Insertions-Varianten zeichnet sich durch die Insertion von Intronsequenzen aus, die zu vorzeitigen Translationsstop führen. Anstelle der normalerweise benutzten 5'-Splice Donorsequenz von Intron 5 wird eine alternative, 3'-lokalisierte Splicestelle in Variante INS1 benutzt, wodurch eine Insertion der ersten 38 bp aus Intron 4 zwischen Exon 4 und Exon 5 entsteht. Ebenso resultiert die

30

Insertion eines Intron 11-Sequenzbereichs in Variante INS2 aus der Benutzung einer alternativen 5'-Splice Donorsequenz in Intron 11. Da diese Variante in der Literatur (Kilian et al., 1997) nur unzureichend beschrieben wurde, läßt sich die genaue alternative 5'-Splice Donorsequenz dieser Variante nicht bestimmen. Die Insertion von Intron 14 Sequenzen zwischen Exon 14 und Exon 15 in Variante INS3 entsteht durch die Benutzung von einer alternativen 3'-Splice Akzeptorsequenz, wodurch der 3'-Teil von Intron 14 nicht gesplitt wird.

Die in unserer Patentanmeldung PCT/EP/98/03469 beschriebene hTC-Variante INS4 (Variante 4) zeichnet sich durch den Ersatz von Exon 15 und dem 5'-Teilbereich von Exon 16 durch die ersten 600 bp des Introns 14 aus. Diese Variante ist auf den Gebrauch einer alternativer internen 5'-Splice Donorsequenz in Intron 14 und einer alternativen 3'-Splice Akzeptorsequenz in Exon 16 zurückzuführen, woraus ein veränderter C-Terminus resultiert.

Die *in vivo*-Generation wahrscheinlich nicht-funktioneller hTC-Proteinvarianten, die mit der Funktion des vollständigen hTC-Proteins interferieren könnten, stellt zusätzlich zur Transkriptionsregulation einen möglichen Mechanismus dar, um die hTC-Proteinfunktion zu kontrollieren. Bis heute ist die Funktion der hTC-Splicevarianten nicht bekannt. Obwohl die meisten dieser Varianten vermutlich für Proteine ohne Reverse-Transkriptase-Aktivität kodieren, könnten sie dennoch eine entscheidende Rolle als transdominant-negative Telomerase-Regulatoren spielen, indem sie z.B. um die Interaktion mit wichtigen Bindungspartnern kompetieren.

Die Suche nach möglichen Transkriptionsfaktorbindungsstellen wurde mit dem „Find Pattern“-Algorithmen aus dem „GCG Sequenz Analysis“ Programmpaket der „Genetics Computer Group“ (Madison, USA) durchgeführt. Dadurch wurden verschiedene potentielle Bindungsstellen für Transkriptionsfaktoren in der Nukleotidsequenz von Intron 2 identifiziert, die in der Tab. 2 aufgelistet sind. Darüberhinaus wurde im Intron 1 eine Sp1-Bindungsstelle (Pos. 43) und im 5'-

untranslatiertem Bereich eine c-Myc-Bindungsstelle (cDNA-Position 29-34, vergl. Fig. 6) gefunden.

Beispiel 6

5

Um den oder die Startpunkt(e) der hTC-Transkription in HL 60 Zellen zu ermitteln, wurde das 5'-Ende der hTC-mRNA durch Primer-Extension-Analyse bestimmt.

Es wurden 2 µg PolyA⁺-RNA aus HL-60-Zellen für 10 min bei 65°C denaturiert. Zur
10 Primeranlagerung wurden 1 µl RNasin (30-40 U/ml) und 0,3-1 pmol radioaktiv
markierter Primer (5'GTTAAGTTGTAGCTTACACTGGTTCTC 3'; 2,5-8x10⁵
cpm) zugegeben und für 30 min bei 37°C in einem Gesamtvolumen von 20 µl
inkubiert. Nach Zugabe von 10 µl 5xReverse Transkriptase-Puffer (Fa. Gibco-BRL),
2 µl 10 mM dNTPs, 2 µl RNasin (s.o.), 5 µl 0,1 M DTT (Fa. Gibco-BRL) 2 µl
15 ThermoScript RT (15 U/µl; Fa. Gibco-BRL) und 9 µl DEPC-behandeltes Wasser
erfolgte die Primer-Verlängerung in einem Gesamtvolumen für 1 h bei 58°C. Die
Reaktion wurde durch 4 µl 0,5 M EDTA, pH 8,0, gestoppt und die RNA nach
Zugabe von 1 µl RNaseA (10 mg/ml) für 30 min bei 37°C abgebaut. Hierauf wurden
2,5 µg gescherte Kalbsthymus-DNA und 100 µl TE addiert und einmal mit 150 µl
20 Phenol/Chloroform (1:1) extrahiert. Die DNA wurde unter Zusatz von 15 µl 3 M Na-
Acetat und 450 µl Ethanol für 45 min bei -70°C gefällt und anschließend für 15 min
bei 14000 Upm abzentrifugiert. Das Präzipitat wurde einmal mit 70 %igem Ethanol
gewaschen, luftgetrocknet und in 8 µl Sequenzierungs-Stopplösung gelöst. Nach 5
min Denaturierung bei 80°C wurden die Proben auf ein 6 %iges Polyacrylamidgel
25 aufgetragen und elektrophoretisch (Ausubel et al., 1987) aufgetrennt (Fig. 5).

Hierbei wurde eine Haupt-Transkriptionsstartstelle identifiziert, die 1767 bp 5' vom
ATG-Startcodon der hTC-cDNA Sequenz lokalisiert ist (Nukleotidposition 3346 in
Fig. 4). Die Nukleotidsequenz um diesen Haupttranskriptionsstart (TTA_nTTGT)
30 repräsentiert darüberhinaus ein Initiator-Element (Inr), das in 6 von 7 Nukleotiden

mit dem Konsensusmotiv (PyPyA_n/Na/tPyPy) (Smale, 1997) eines Initiator-Elementes übereinstimmt.

5 In unmittelbarer Nähe des experimentell identifizierten Haupt-Transkriptionsstartes konnte keine eindeutige TATA-Box identifiziert werden, so daß der hTC-Promoter wahrscheinlich in die Familie der TATA-losten Promotoren (Smale, 1997) einzuordnen ist. Allerdings wurde durch Bioinformatik Analyse eine potentielle TATA-Box von Nukleotidposition 1306 bis 1311 (Fig. 4) gefunden. Die zusätzlich um den Haupt-Transkriptionsstart beobachteten Neben-Transkriptionsstarts wurden auch bei
10 anderen TATA-losten Promotoren beschrieben (Geng and Johnson, 1993), wie z.B. in den stark regulierten Promotoren einiger Zellzyklusgene (Wick *et al.*, 1995).

Beispiel 7

15 Zusätzlich zu dem in Beispiel 6 beschriebenen, in HL60 Zellen identifizierten Startpunkt des hTC Transkriptes, wurde ein weiterer Transkriptionsstartbereich in HL60 Zellen identifiziert. Anhand von RT-PCR-Analysen wurde die Region des Transkriptionsstarts des hTC-Gens in HL60 Zellen auf die bp -60 bis -105 eingegrenzt.

20 Unter Einsatz von 0,4 µg Poly A-RNA aus HL60 Zellen (Clontech) und dem genspezifischen Primer GSP13 (5'-CCTCCAAAGAGGTGGCTTCTTCGGC-3', cDNA-Position 920-897) wurde hierfür die cDNA mit Hilfe des „First Strand cDNA-Synthesis Kit“ (Clontech) nach Angaben der Hersteller synthetisiert. In einem
25 Endvolumen von 50 µl wurden 1 µl cDNA mit 10 pmol dNTP-Mix versetzt und in 1xPCR-Reaktionspuffer F (PCR-Optimizer Kit der Fa. InVitrogen) und einem Unit Platinum-Taq-DNA Polymerase (Fa. Gibco/BRL) eine PCR-Reaktion durchgeführt. Als Primer wurden jeweils 10 pmol der nachfolgend definierten 5'- und 3'-Primer
30 zugefügt. Die PCR wurde in 3 Schritten durchgeführt. An eine zweiminütige Denaturierung bei 94°C schlossen sich 36 PCR-Zyklen an, in denen die DNA zunächst für 45 sec bei 94°C denaturiert wurde und anschließend für 5 min bei 68°C die Primer

angelagert und die DNA-Kette verlängert wurde. Zum Abschluß folgte für 10 min eine Kettenverlängerung bei 68°C. Insgesamt wurden sechs verschiedene 5'-PCR Primer (Primer HTRT5B: 5'-CGCAGCCACTACCGCGAGGTGC-3', cDNA-Position 105 bis 126; Primer C5S: 5'-CTGCGTCCTGCTGCGCACGTGGGAAGC-3', 5'-flankierende Region -49 bis -23; Primer PRO-TEST1: 5'-CTCGCGGCGCGAGTTTCAGGCAG-3', 5'-flankierende Region -74 bis -52; Primer PRO-TEST2: 5'-CCAGCCCCTCCCCTTCCTTTCC-3', 5'-flankierende Region -112 bis -91; Primer PRO-TEST4: 5'-CCAGCTCCGCCTCCTCCGCGC-3', 5'-flankierende Region -191 - -171; Primer RP-3A: 5'-CTAGGCCGATTCGACCTCTCTCC-3', 5'-flankierende Region -427 bis -405) mit dem 3'-PCR Primer C5Rück (5'-GTCCCAGGGCAGCACACCAG-3', cDNA-Position 245 bis 225) kombiniert. Als Kontrolle wurde zusätzlich zu den Oligo-dT- und GSP13-geprimten cDNAs auch genomische DNA für die PCR eingesetzt. Wie in Fig. 9 gezeigt, wurde nur mit den Primerkombinationen HTRT5B-C5Rück, C5S-C5Rück und PRO-TEST1-C5Rück ein PCR-Produkt erhalten, was darauf hinweist, daß der Startpunkt der hTC-Transkription in der Region zwischen bp-60 und bp-105 liegt.

Beispiel 8

In der ca. 11,2 kb isolierten 5'-flankierenden Region des hTC-Gens befinden sich mehrere extrem GC-reiche Bereiche, sog. CpG Islands. Ein CpG Islands mit einem GC-Gehalt von > 70 % reicht von bp - 1214 bis in Intron 2. Zwei weitere GC-reiche Bereiche mit einem GC-Gehalt von > 60 % reichen von bp -3872 bis bp -3113 bzw. bp -5363 bis bp -3941. Die Lage der CpG Islands ist in der Fig. 11 graphisch dargestellt.

Die Suche nach möglichen Transkriptionsfaktorbindungsstellen wurde mit dem „Find Pattern“-Algorithmus aus dem „GCG Sequenz Analysis“ Programmpaket der „Genetics Computer Group“ (Madison, USA) durchgeführt. Dadurch wurden verschiedene potentielle Bindungsstellen in der Region bis -900 bp upstream vom

Translations-Startcodon ATG indentifiziert: fünf Sp1-Bindungsstellen, eine c-Myc-Bindungsstelle, eine CCAC-Box (Fig. 10). Zusätzlich wurden eine CCAAT-Box und eine zweite c-Myc-Bindungsstelle an den Positionen -1788 bzw. -3995 der 5'-flankierenden Region gefunden.

5

Beispiel 9

Um die Aktivität des hTC-Promotors zu analysieren, wurden durch PCR-Amplifikation vier verschieden lange hTC-Promotorsequenzabschnitte generiert und 5' vor das Reportergen Luziferase in den Vektor pGL2 der Fa. Promega kloniert. Als DNA-Quelle für die PCR-Amplifikation wurde das aus dem Phagenklon P12 subklonierte, 8,5 kb große SacI-Fragment gewählt. In einem Endvolumen von 50 µl wurden 35 ng dieser DNA mit 10 pmol dNTP-Mix versetzt und in 1xPCR-Reaktionspuffer (PCR-Optimizer Kit der Fa. InVitrogen) und einem Unit Platinum-Taq-DNA Polymerase (Fa. Gibco/BRL) eine PCR-Reaktion durchgeführt. Als Primer wurden jeweils 20 pmol der nachfolgend definierten 5'- und 3'-Primer zugefügt. Die PCR wurde in 3 Schritten durchgeführt. An eine zweiminütige Denaturierung bei 94°C schlossen sich 30 PCR-Zyklen an, in denen die DNA zunächst für 45 sec bei 94°C denaturiert wurde und anschließend für 5 min bei 68°C die Primer angelagert und die DNA-Kette verlängert wurde. Zum Abschluß folgte für 10 min eine Kettenverlängerung bei 68°C. Als 3'-PCR-Primer wurde jeweils der Primer PK-3A (5'-GCAAGCTTGACGCAGCGCTGCCTGAAACTCG-3', Position -43 bis -65) gewählt, der einen Sequenzbereich 42 bp upstream vom START-Codon ATG erkennt. Durch Kombination des PK-3A-Primers mit dem 5'-PCR-Primer PK-5B (5'-CCAGATCTCTGGAACACAGAGTGGCAGTTTCC-3', Position -4093 bis -4070) wurde ein 4051 bp großes Promotor-Fragment amplifiziert (NPK8). Die Kombination des Primerpaares PK-3A und PK-5C (5'-CCAGATCTGCATGAAGTGTGTGGGATTTCAG-3', Position -3120 bis -3096) führte zur Amplifikation eines 3078 bp großen Promotorfragmentes (NPK15). Ein 2068 bp großes Promotorfragment wurde durch die Verwendung der Primerkombination PK-3A und PK-5D (5'-

GGAGATCTGATCTTGGCTTACTGCAGCCTCTG-3', Position -2110 bis -2087) amplifiziert (NPK22). Der Einsatz der Primerkombination PK-3A und PK-5E (5'-GGAGATCTGTCTGGATTCCTGGGAAGTCCTCA-3', Position -1125 bis -1102) führte schließlich zur Amplifikation eines 1083 bp großen Promotorfragmentes (NPK27). Der PK-3A Primer enthält eine HindIII-Erkennungssequenz. Die
5 verschiedenen 5'-Primer enthalten eine BglII-Erkennungssequenz.

Die entstandenen PCR-Produkte wurden mit Hilfe des QIA quick spin PCR Purification Kits der Fa. Qiagen nach Angaben der Hersteller aufgereinigt und
10 anschließend mit den Restriktionsenzymen BglII und HindIII verdaut. Mit den gleichen Restriktionsenzymen wurde der pGL2-Promotor-Vektor verdaut und der in diesem Vektor enthaltene SV40-Promotor freigesetzt und abgetrennt. Die PCR-Promotorfragmente wurden in den Vektor ligiert, in kompetente DH5 α -Bakterien der Fa. Gibco/BRL transformiert. Aus transformierten Bakterienklonen wurde DNA für
15 die nachfolgend beschriebenen Promotor-Aktivitäts-Analysen mit Hilfe des Qiagen Plasmid-Kits der Fa. Qiagen isoliert.

Beispiel 10

20 Die Aktivität des hTC-Promotors wurde in transienten Transfektionen in eukaryotischen Zellen analysiert.

Alle Arbeiten mit eukaryotischen Zellen erfolgten an einem sterilen Arbeitsplatz. CHO-K1 und HEK 293 Zellen wurden von der American Type Culture collection
25 bezogen.

CHO-K1 Zellen wurden in DMEM Nut Mix F-12 Zellkulturmedium (Fa. Gibco-BRL, Bestellnummer: 21331-020) mit 0,15 % Streptomycin/Penezillin, 2 mM Glutamin und 10 % FCS (Fa. Gibco-BRL) gehalten.

30

HEK 293 Zellen wurden in DMOD Zellkulturmedium (Fa. Gibco-BRL, Bestellnummer: 41965-039) mit 0,15 % Streptomycin/Penizillin, 2 mM Glutamin und 10 % FCS (Fa. Gibco-BRL) kultiviert.

- 5 CHO-K1 und HEK 293 Zellen wurden in wassergesättigter Atmosphäre bei 37°C unter Begasung mit 5 % CO₂ kultiviert. Bei konfluentem Zellrasen wurde das Medium abgesaugt, die Zellen mit PBS (100 mM KH₂PO₄, pH 7,2; 150 mM NaCl) gewaschen und durch Zugabe einer Trypsin-EDTA Lösung (Fa. Gibco-BRL) abgelöst. Das Trypsin wurde durch Mediumzugabe inaktiviert und die Zellzahl mit
10 einer Neubauer-Zählkammer ermittelt, um die Zellen in gewünschter Dichte auszu-plattieren.

- Für die Transfektion wurden pro Well jeweils 2×10^5 HEK 293 Zellen in einer 24-well Zellkulturplatte ausplattiert. Nach 3 Stunden wurde das HEK 293 Medium
15 entfernt. Für die Transfektion wurden bis zu 2,5 µg Plasmid-DNA, 1 µg eines CMV β-Gal Plasmidkonstruktes (Fa. Stratagene, Bestellnummer: 200388), 200 µl serum-freies Medium und 10 µl Transfektionsreagenz (DOTAP der Fa. Boehringer Mannheim) für 15 Minuten bei Raumtemperatur inkubiert und anschließend auf die HEK 293 Zellen gleichmäßig aufgetropft. Nach 3 Stunden wurden 1,5 ml Medium
20 hinzugegeben. Nach 20 Stunden wurde das Medium gewechselt. Nach weiteren 24 Stunden wurden die Zellen zur Bestimmung der Luziferase- und der β-Gal-Aktivität geerntet. Dazu wurden die Zellen im Zellkultur-Lysisreagenz (25 mM Tris [pH 7,8] mit H₃PO₄; 2 mM CDTA; 2 mM DTT; 10% Glycerol; 1% Triton X-100) für 15 Minuten bei Raumtemperatur lysiert. Zwanzig µl dieses Zellysats wurden mit 100 µl
25 Luziferase-Assaypuffer (20 mM Tricin; 1,07 mM (MgCO₃)₄ Mg(OH)₂·5H₂O; 2,67 mM MgSO₄; 0,1 mM EDTA; 33,3 mM DTT; 270 µM Coenzym A; 470 µM Luciferin, 530 µM ATP) gemischt und das durch die Luziferase generierte Licht gemessen.

- 30 Zur Messung der β-Galaktosidaseaktivität wurden gleiche Mengen Zellysats und β-Galaktosidase-Assaypuffer (100 mM Natriumphosphatpuffer pH 7,3; 1 mM MgCl₂;

50 mM β -Merkaptoethanol; 0,665 mg/ml ONPG) für mindestens 30 Minuten bei 37°C oder bis eine leichte Gelbfärbung auftrat, inkubiert. Die Reaktion wurde durch Zugabe von 100 μ l 1 M Na_2CO_3 gestoppt und die Absorption bei 420 nm bestimmt.

- 5 Für die Analyse des hTC-Promotors wurden vier verschieden lange hTC-Promotorsequenzabschnitte 5' vor das Reportergen Luziferase kloniert (vergl. Beispiel 9).

In der Fig. 11 sind die relativen Luziferase Aktivitäten zweier unabhängiger Transfektionen mit den Konstrukten NPK8, NPK15, NPK22 und NPK27 in HEK
10 293 Zellen aufgetragen. Jedes Experiment wurde in Duplikaten durchgeführt. Darüberhinaus wurde die Standardabweichung angegeben. Das Konstrukt NPK 27 zeigt eine 40fach höhere Luziferaseaktivität als die Basalaktivität des promotorlosen Luziferase-Kontrollkonstrutes (pGL2-basic) und eine 2 bis 3fach höhere Aktivität als das SV40 Promotorkontroll-Konstrukt (pGL2PRO). Interessanterweise wurde im
15 Vergleich zu dem Konstrukt NPK27 eine 2 bis 3fach geringere Luziferaseaktivität in mit längeren hTC Promotorkonstrukten (NPK8, NPK15, NPK22) transfizierten Zellen beobachtet. Ähnliche Ergebnisse wurden auch in CHO Zellen beobachtet (Daten nicht gezeigt).

Literaturverzeichnis

- 5 Allsopp, R. C., Vazire, H., Patterson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W. und Harley, C.B. (1992). Telomere length predicts replicative capacity of human fibroblasts. *Proc. Natl. Acad. Sci.* 89, 10114-10118.
- 10 Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1987). *Current protocols in molecular biology*. Greene Publishing Associates and Wiley-Intersciences, New York.
- Blasco, M. A., Rizen, M., Greider, C. W. und Hanahan, D. (1996). Differential regulation of telomerase activity and telomerase RNA during multistage tumorigenesis. *Nature Genetics* 12, 200-204.
- 15 Broccoli, D., Young, J. W. und deLange, T. (1995). Telomerase activity in normal and malignant hematopoietic cells. *Proc. Natl. Acad. Sci.* 92, 9082-9086.
- Counter, C. M., Avilion, A. A., LeFeuvre, C. E., Stewart, N. G. Greider, C.W. Harley, C. B. und Bacchetti S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* 11, 1921-1929.
- 20 Feng, J., Funk, W. D., Wang, S.-S., Weinrich, S. L., Avilion, A.A., Chiu, C.-P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., Le, S., West, M.D., Harley, C.B., Andrews, W.H., Greider, C.W. und Villeponteau, B. (1995). The RNA component of human telomerase. *Science* 269, 1236-1241.
- 25 Geng, Y., and Johnson, L.F. (1993). Lack of an initiator element is responsible for multiple transcriptional initiation sites of the TATA less mouse thymidine synthase promoter. *Mol. Cell. Biol.* 14:4894.
- 30 Goldstein, S. (1990). Replicative senescence: The human fibroblast comes of age. *Science* 249, 1129-1133.
- 35 Harley, C.B., Futcher, A.B., Greider, C.W., 1990. Telomeres shorten during ageing of human fibroblasts. *Nature* 345, 458-460.

- Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D.K. und Allshire, R.C. (1990). Telomere reduction in human colorectal carcinoma and with ageing. *Nature* 346, 866-868.
- 5 Hiyama, K., Hirai, Y., Kyoizumi, S., Akiyama, M., Hiyama, E., Piatyszek, M.A., Shay, J.W., Ishioka, S. und Yamakido, M. (1995). Activation of telomerase in human lymphocytes and hematopoietic progenitor cells. *J. Immunol.* 155, 3711-3715.
- 10 Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C. B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L. und Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* 266, 2011-2015.
- Latchman, D.S. (1991). Eukaryotic transcription factors. Academic Press Limited, London.
- 15 Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V. und Cech T.R. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* 276: 561-567.
- Lundblad, V. und Szostak, J. W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57, 633-643.
- 20 McClintock, B. (1941). The stability of broken ends of chromosomes in *Zea mays*. *Genetics* 26, 234-282.
- 25 Meyne, J., Ratliff, R. L. und Moyzis, R. K. (1989). Conservation of the human telomere sequence (TTAGGG)_n among vertebrates. *Proc. Natl. Acad. Sci.* 86, 7049-7053.
- Olovnikov, A. M. (1973). A theory of marginotomy. *J. Theor. Biol.* 41, 181-190.
- 30 Sandell, L. L. und Zakian, V. A. (1993). Loss of a yeast telomere: Arrest, recovery and chromosome loss. *Cell* 75, 729-739.
- Shapiro, M.B., Senapathy, P., 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucl. Acids Res.* 15, 7155-7174.
- 35 Smale, S.T. and Baltimore, D. (1989). The „initiator“ as a transcription control element. *Cell* 57:103-113.

- Smale, S.T. (1997). Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochimica et Biophysica Acta* 1351, 73-88.
- 5 Shay, J. W. (1997). Telomerase and Cancer. Ciba Foundation Meeting: Telomeres and Telomerase. London.
- Vaziri, H., Dragowska, W., Allsopp, R. C., Thomas, T. E., Harley, C.B. und Landsdorp, P.M. (1994). Evidence for a mitotic clock in human hematopoietic stem cells: Loss of telomeric DNA with age. *Proc. Natl. Acad. Sci.* 91, 9857-9860.
- 10 Wick, M., Härönen, R., Mumberg, D., Bürger, C., Olsen, B.R., Budarf, M.L., Apte, S. S. and Müller, R. (1995). Structure of the human TIMP-3 gene and its cell-cycle-regulated promoter. *Biochemical Journal* 311, 549-554.
- 15 Zakian, V. A. (1995). Telomeres: Beginning to understand the end. *Science* 270, 1601-1607.

Patentansprüche

1. Regulatorische DNA-Sequenzen für das Gen der humanen katalytischen
Telomerase-Untereinheit.
- 5 2. DNA-Sequenzen gemäß Anspruch 1, dadurch gekennzeichnet, daß es sich um
Intronsequenzen gemäß SEQ ID NO 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,
17, 18, 19 und/oder 20 oder um regulatorisch wirksame Fragmente dieser
Sequenzen handelt.
- 10 3. DNA-Sequenzen gemäß Anspruch 1, dadurch gekennzeichnet, daß es sich um
die 5'-flankierende regulatorische DNA-Sequenz für das Gen der humanen
katalytischen Telomerase-Untereinheit gemäß Fig. 10 (SEQ ID NO 3) oder
um regulatorisch wirksame Fragmente dieser DNA-Sequenz handelt.
- 15 4. Rekombinantes Konstrukt, enthaltend eine DNA-Sequenz gemäß einem der
Ansprüche 1 bis 3.
5. Rekombinantes Konstrukt gemäß Anspruch 4, dadurch gekennzeichnet, daß
20 es weiterhin eine oder mehrere DNA-Sequenzen enthält, die für Polypeptide
oder Proteine kodieren.
6. Vektor, enthaltend ein rekombinantes Konstrukt gemäß Anspruch 4 oder 5.
- 25 7. Verwendung von rekombinanten Konstrukten bzw. Vektoren gemäß einem
der Ansprüche 4 bis 6 zur Herstellung von Arzneimitteln.
8. Rekombinante Wirtszellen, enthaltend rekombinante Konstrukte bzw. Vek-
toren gemäß einem der Ansprüche 4 bis 6.

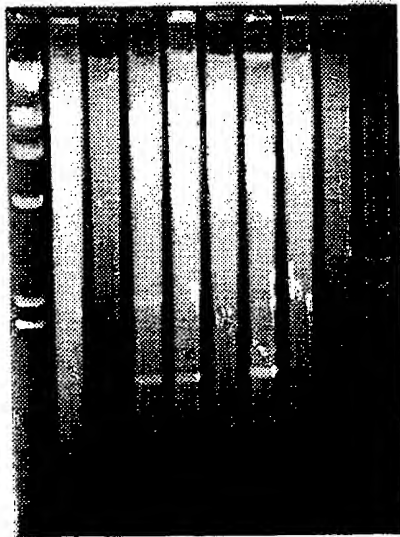
9. Verfahren zur Identifizierung von Substanzen, die die Promotor-, Silencer- oder Enhanceraktivität der humanen katalytischen Telomerase-Untereinheit beeinflussen, das folgende Schritte umfaßt:
- 5 A. Zugabe einer Kandidatensubstanz zu einer Wirtszelle, enthaltend DNA-Sequenzen gemäß einem der Ansprüche 1 bis 3, funktionell verknüpft mit einem Reportergen,
- B. Messung des Substanzeffektes auf die Reportergenexpression.
- 10
10. Verfahren zur Identifizierung von Faktoren, die spezifisch an die DNA gemäß einem der Ansprüche 1 bis 3 oder an Fragmente davon binden, dadurch gekennzeichnet, daß man eine Expressions-cDNA-Bibliothek mit einer DNA-Sequenz gemäß einem der Ansprüche 1 bis 3 oder Teilfragmenten unterschiedlichster Länge als Sonde screent.
- 15
11. Transgene Tiere, enthaltend rekombinante Konstrukte bzw. Vektoren gemäß Ansprüchen 4 bis 6.
- 20
12. Verfahren zur Detektion Telomerase-assoziiierter Zustände bei einem Patienten, das folgende Schritte umfaßt:
- A. Inkubation eines rekombinanten Konstruktes bzw. Vektors gemäß Ansprüchen 4 bis 6 das bzw. der zusätzlich ein Reportergen enthält
- 25 mit Körperflüssigkeiten oder zellulären Proben,
- B. Detektion der Reportergenaktivität, um einen diagnostischen Wert zu erhalten,

- C. Vergleich des diagnostischen Wertes mit Standardwerten für das Reportergenkonstrukt in standardisierten normalen Zellen oder Körperflüssigkeiten des gleichen Typs wie die Testprobe.

Fig. 1

A

1 2 3 4 5 6 7 8 9 10



B

1 2 3 4 5 6 7 8 9 10

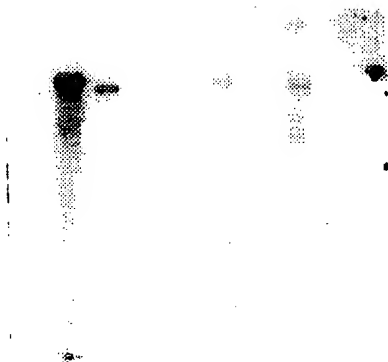
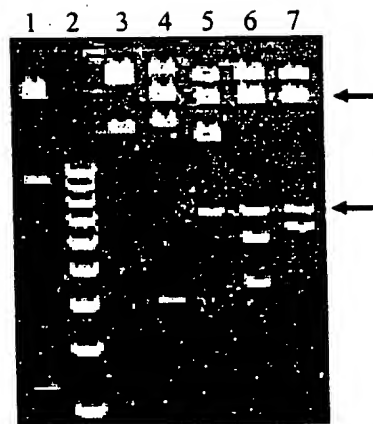


Fig. 2



ERSATZBLATT (REGEL 26)

Fig. 3

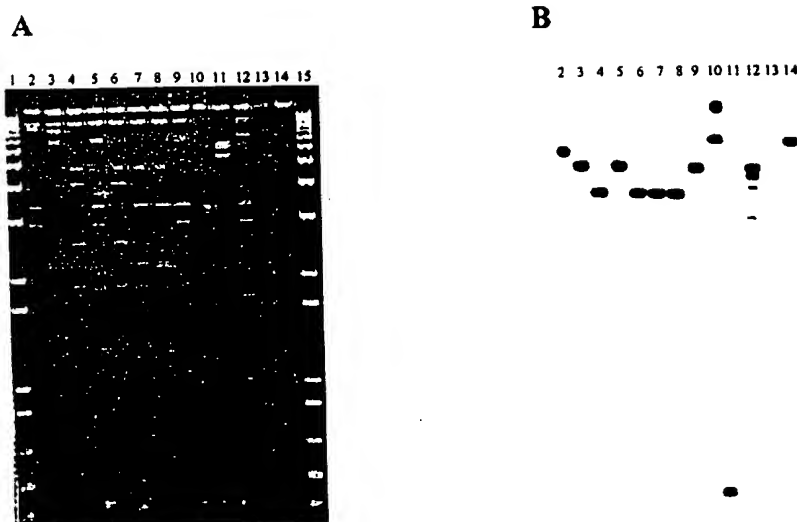


Fig. 4

GAGCTCTGAA CCGTGGAAC GAACATGACC CTTGCTGCC TGCTTCCCTG GGTGGGTCAA GGGTAATGAA 70
 GTGGTGTGCA GGAAATGGCC ATGTAAATTA CAGCACTCTG CTGATGGGGA CCGTTCCTTC CATCATATT 140
 CATCTTCACC CCCAAGGACT GAATGATTCC AGCAACTTCT TCGGGTGTGA CAAGCCATGA CAAAACCTCAG 210
 TACAACACCC ACTCTTTTAC TAGGCCACCA GAGCAGGGGC CACACCCCTG ATATATTAAG AGTCCAGGAG 280
 AGATGAGGCT GCTTTCAGCC ACCAGGCTGG GGTGACAACA GCGGCTGAAC AGTCTGTTC TCTAGACTAG 350
 TAGACCTGG CAGGCACTCC CCCAAATCT AGGGCCTGGT TGCTGCTTCC CGAGGGCGCC ATCTGCCCTG 420
 GAGACTCAGC CTGGGGTGCC ACACTGAGGC CAGCCCTGTC TCCACACCT CCGCCTCCAG GCCTCAGCTT 490
 CTCCAGCAGC TTCTTAAACC CTGGGTGGGC CGTGTTCAG CGCTACTGTC TCACCTGTCC CACTGTGTCT 560
 TGCTCTAGCG ACGTAGCTCG CACGGTTCCT CCTCACATGG GGTGTCTGTC TCCTTCCCCA ACACTCACAT 630
 GCGTTGAAGG GAGGAGATTG TGGGCTCCC AGACTGGCTC CTCTGAGCCT GAACCTGGCT CGTGGCCCCC 700
 GATGCAGGTT CCTGGCGTCC GGCTGCAGGC TGACCTCCAT TTCCAGGCGC TCCCGCTCTC CTGTCACTTG 770
 CCGGGGCTCG CCGGTGTGTT CTCTGTCTT TGTGCTCCTT TCCACGTCCA GCTGCGTGTG TCTCTGCCCG 840
 CTAGGGTCTC GGGGTTTTTA TAGGCATAGG ACGGGGGCGT GGTGGGCCAG GCGCTCTTG GGAATGCAA 910
 CATTTGGGTG TGAAGTAGG AGTGCTGTC CTCACCTAGG TCCACGGGCA CAGGCTGGG GATGGAGCCC 980
 CCGCCAGGGA CCCGCCCTTC TCTGCCAGC ACTTCTCTGC CCCCCTCCCT CTGGAACACA GAGTGGCAGT 1050
 TTCCACAAGC ACTAAGCATC CTCTTCCCA AAGACCCAGC ATTGGCAGCC CTGGACATT GCCCCACAGC 1120
 CCTGGGAATT CAGGTGACTA CGCACATCAT GTACACACT CCGTCCACGA CCGACCCCG CTGTTTTATT 1190
 TTAATAGCTA CAAAGCAGGG AAATCCCTGC TAAATGTCC TTTAACAAAC TGGTTAAACA AACGGGTCCA 1260
 TCCGCACGGT GGACAGTTCC TCACAGTGAA GAGGAACATG CCGTTTATAA AGCCTGCAGG CATCTCAAGG 1330
 GAATTACGCT GAGTCAAAAC TGCCACTCC ATGGGATACG TACGCAACAT GCTCAAAAG AAGAATTTTC 1400
 ACCCATGGC AGGGGAGTGG TTAGGGGGGT TAAGGACGGT GGGGGCGGCA GCTGGGGCT ACTGCACGCA 1470
 CCTTTTACTA AAGCCAGTTT CTTGGTCTG ATGGTATTGG CTCAGTTATG GGAGACTAAC CATAGGGGAG 1540
 TGGGGATGGG GGAACCCGGA GGCTGTGCCA TCTTGCCAT GCCCGAGTGT CCGGGCAGG ATAATGCTCT 1610
 AGAGATGCCC ACGTCTGAT TCCCCAAAC CTGTGGACAG AACCCGCCCG GCCCCAGGGC CTTTGCAGGT 1680
 GTGATCTCCG TGAGGACCTT GAGGTCTGGG ATCCTTCGGG ACTACCTGCA GGGCCGAAAA GTAATCCAGG 1750
 GGTCTCTGGA AGAGGGCGGC AGGAGGTCA GAGGGGGGCA GCCTCAGGAC GATGGAGGCA GTCAGTCTGA 1820
 GGCTGAAAG GAGGGGAGGG CCTCGAGGCC AGGCCTGCAA CCGCCTCCAG AAGCTGGAAG AAGCGGGGAA 1890
 GGGACCTCC ACGGAGCCTG CAGCAGGAAG GCACGGCTGG CCCTTAGCCC ACCAGGGCCC ATCGTGGACC 1960
 TCCGGCTCC GTGCCATAGG AGGGCACTCG CGCTGCCCTT CTAGCATGAA GTGTGTGGGG ATTTGCAGAA 2030
 GCAACAGGAA ACCCATGCAC TGTGAATCTA GATTATTTT GGCAGGCAG AGTGATTTTA TTTAGCTATT TATTTTATT 2100
 CAGGGCTGAA GTGCCCTCG GCAAGGCGAG GGCAGGCAG AGTGATTTTA TTTAGCTATT TATTTTATT 2170
 TACTTACTTT CTGAGACAGA GTTATGCTCT TGTGGCCAG GCTGGAGTGC AGCGGCATGA TCTTGGCTCA 2240
 CTGCAACCTC CGTCTCTCG GTTCAAGCAA TTCTCGTGCC TCAGCCTCCC AAGTAGCTGG GATTTCAGGC 2310
 GTGACCCACC ACACCCGCT AATTTTGAT TTTTAGTAGA GATGGGCTTT CACCATGTTG GTCAAGCTGA 2380
 TCTCAAAATC CTGACCTCAG GTGATCCGCC CACCTCAGCC TCCCAAAGTG CTGGGATTAC AGGCATGAGC 2450
 CACTGCACCT GGCCTATTTA ACCATTTTAA AACTTCCCTG GGCTCAAGTC ACACCCACTG GTAAGGAGTT 2520
 CATGGAGTTC AATTCTCCCT TTACTCAGGA GTTACCCTCC TTTGATATT TCTGTAAATC TTCGTAGACT 2590
 GGGGATACAC CGTCTCTTGA CATATTACA GTTCTGTGA CCACCTGTTA TCCCATGGGA CCCACTGCAG 2660
 GGGCAGCTGG GAGGCTGCAG GCTTCAGGTC CCAGTGGGGT TGCCATCTGC CAGTAGAAAC CTGATGTAGA 2730
 ATCAGGGCGC AAGTGTAGAC ACTGTCCTGA ATCTCAATGT CTCAGTGTGT GCTGAAACAT GTAGAAATTA 2800
 AAGTCCATCC CTCCTACTCT ACTGGGATTG AGCCCTTCC CTATCCCCC CCAGGGGCG AGGAGTTCTT 2870
 CTCACTCTG TGGAGGAAGG AATGATATT TGTATTTTT CACTGCTGGT ACTGAATCCA CTGTTTCATT 2940
 TGTGTGTTG TTTGTTTTGT TTTGAGAGGC GGTTCACCTC TGTGTGCTCA GGTGGAGGG AGTGCATGG 3010
 CCGGATCTTG GCTTACTGCA GCCTCTGCCT CCCAGGTTCA AGTGATTCTC CTGCTTCCCG CTCCCATTTG 3080
 GCTGGGATTA CAGGCACCCG CCACCATGCC CAGCTAAATT TTTGATTTT TAGTAGAGAC GGGGGTGGGT 3150
 GGGGTTCAAC ATGTTGGCCA GGCTGGTCTC GAACCTCTGA CCTCAGATGA TCCACCTGCC TCTGCTCTCT 3220
 AAAGTGCTGG GATTACAGGT GTGAGCCACC ATGCCAGCT CAGAATTAC TCTGTTTAGA AACATCTGGG 3290
 TCTGAGGTAG GAAGCTCACC CCACTCAAGT GTTGTGGTGT TTTAAGCCAA TGATAGAAAT TTTTATTGT 3360
 TGTTAGAACA CTCTTGATGT TTTACTGTGT GATGACTAAG ACATCATCAG CTTTTCAAAG ACACACTAAC 3430
 TGCACCCATA AACTTGGGGT GTCTTCTGGG TATCAGCAAT CTTCATTGAA TGCCGGGAGG CGTTTCTCTG 3500
 CCATGCACAT GGTGTAAAT ACTCCAGCAT AATCTTCTGC TTCCATTCT TCTCTTCCCT CTTTAAAT 3570
 TGTGTTTTCT ATGTTGGCTT CTCTGCAGAG AACCAGTGA AGCTACAAT TAACTTTTGT TGAACAAAT 3640
 TTTCCAAACC GCCCTTTGCT CTAAGTGCCA GAGACAATC ACAAACACAG CCGTTTAAAG AGGCTTAGGG 3710
 ATCACTAAGG GGATTCTAG AAGAGCGACC TGTAATCTTA AGTATTTACA AGACGAGGCT AACCTCCAGC 3780
 GAGCGTGACA GCCCAGGGAG GGTGCGAGGC CTGTTCAAAAT GCTAGCTCCA TAAATAAAGC AATTTCCTCC 3850
 GGCAGTTTCT GAAAGTAGGA AAGGTTACAT TTAAGGTTGC GTTGTGTTAGC ATTTCAAGTG TTGCCGACCT 3920
 CAGCTACAGC ATCCCTGCAA GGCCTCGGGA GACCCAGAAG TTTCTCGCCC CCTTAGATCC AAACCTGAGC 3990
 AACCCGGAGT CTGGATTCTT GGGAAAGTCT CAGCTGTCTT GCGGTGTGTC CCGGGCCGCA GGTCTGGAGG 4060
 GGACCACTGG CCGTGTGGCT TCTACTGCTG GGCTGGAAGT CCGGCTCTCAG AGCTCTGAG TCCGAGGCTT 4130
 GGAGCCAGGT GCCTGGACCC CGAGGCTGCC CTCCACCTG TGCGGGCGGG ATGTGACCAG ATGTTGGCCT 4200
 CATCTGCCAG ACAGAGTGCC GGGGCCAGG GTCAAGCCCG TTGTGGCTGG TGTGAGGGCG CCGGTGGCGG 4270
 GTCAGCAGGA CGGCTGGCT CCATTTCCCA CCGTTTCTG ACGGGACCGC CCGGTGGGT GATTACAGA 4340
 TTTGGGGTGG TTTGCTCATG GTGGGGACCC CTCGCCGCTC GAGAACCTGC AAAGAGAAAT GACGGGCTG 4410
 TGTCAAGGAG CCAAGTCCG GGGGAAGTGT TGCAGGAGG CACTCCGGGA GGTCCCGCTT GCCCGTCCAG 4480
 GGAGCAATGC GTCTCGGGT TCGTCCCGAG CCGCGTCTAC GCGCCTCCGT CCTCCCTTC ACGTCCGGCA 4550
 TCGTGGTGC CCGGAGCCCG ACGCCCCGCG TCCGAGCTGC GAGGCAGCCC TGGGTCTCCG GATCAGGCCA 4620
 CGGGCCAAAG GGTGCCCGCA CGCACTGTT CCCAGGGCCT CCACATCATG GCCCTCCCT CCGGTACCC 4690

Fig. 4 (Fortsetzung)

```

CACAGCCTAG GCCGATTCTGA CCTCTCTCCG CTGGGGCCCT CGCTGGCGTC CCTGCACCCT GGGAGCGCGA 4760
GCGGCGCGCG GCGGGGGAAG CGCGGCCAG ACCCCCGGGT CCGCCCGGAG CAGCTGCGCT GTCGGGGCCA 4830
GGCCGGGCTC CCAGTGGATT CGCGGGCACA GACGCCAGG ACCGCGCTCC CCACGTGGCG GAGGGACTGG 4900
GGACCCGGGC ACCCGTCTG CCCCTTCACC TTCCAGCTCC GCCTCCTCCG CGCGGACCCC GCCCCGTCCC 4970
GACCCCTCCC GGGTCCCCGG CCCAGCCCCC TCCGGGCCCT CCCAGCCCCT CCCCTTCCTT TCCGCGGGCC 5040
CGCCCTCTCC TCGCGGCGCG AGTTTCAGGC AGCGCTGCGT CCTGCTGCGC ACGTGGGAAG CCCTGGCCCC 5110
GGCCACCCCC GCGATG
5126

```

Fig. 5

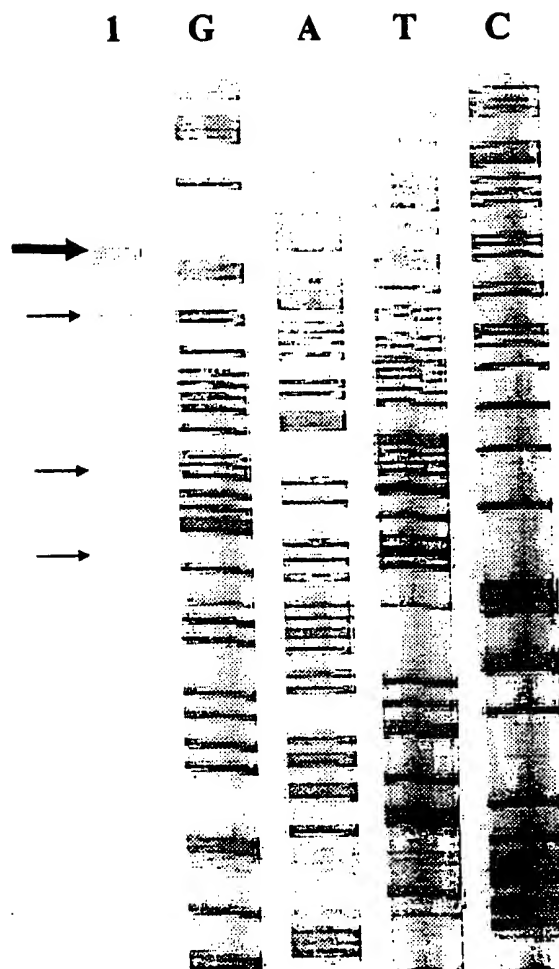


Fig. 6

GTTTCAGGCA GCGCTGCGTC CTGCTGCGCA CGTGGGAAGC CCTGGCCCCG GCCACCCCCG CGATGCCCGG 70
 CGCTCCCCCG TGCCGAGCCG TGCGCTCCCT GCTGCGCAGC CACTACCCGG AGGTGCTGCC GCTGGCCACG 140
 TTCGTGCGGC GCCTGGGGCC CCAGGGCTGG CGGCTGGTGC AGCGCGGGGA CCCGGCGGCT TTCGCGCGC 210
 TGGTGGCCCA GTGCTGGTGT TGGCTGCCCT GGGACGCACG GCGCGCCCCC GCGCGCCCTT CCTTCCGCCA 280
 GGTGTCTCTG CTGAAGGAGC TGGTGGCCCC AGTGTCTCAG AGGCTGTGCG AGCGCGGCGC GAAGAAGCTG 350
 CTGGCCCTTC GCTTCGCGCT GCTGGACGGG GCGCGCGGGG GCGCGCCCCG GGCCTTCACC ACCAGCGTGC 420
 GCAGCTACCT GCCAACACG GTGACCGACG CACTGCGGGG GAGCGGGGCG TGGGGGCTGC TGTGCGCCG 490
 CGTGGGCGAC GACGTGCTGG TTCACCTGCT GGCACGCTGC GCGCTCTTTG TGTGGTGGC TCCAGAGTGC 560
 GCCTACCAGG TGTGCGGGCC GCGGCTGTAC CAGCTCGGCG CTGCCACTCA GCGCGGCCCC CCGCCACACG 630
 CTAGTGGACC CCGAAGGCGT CTGGGATGCG AACGGGCTG GAACCATAGC GTCAGGGAGG CCGGGGTCCC 700
 CCTGGGCGTG CCAGCCCCGG GTGCGAGGAG GCGCGGGGGC AGTGCCAGCC GAAGTCTGCC GTTGCCCAAG 770
 AGGCCACGGC GTGGCGCTGC CCCTGAGCGG GAGCGGACGC CCGTTGGGCA GGGGTCTCTG GCCCACCCGG 840
 GCAGGACGGC TGGACCGAGT GACCGTGGTT TCTGTGTGGT GTCACCTGCC AGACCCGCGC AAGAAGCCAC 910
 CTCTTTGGAG GGTGCGCTCT CTGGCACGCG CCACTCCAC CCATCCGTGG CCGCCACAGC CCACGCGGGC 980
 CCCCACCCA CATCGCGGCC ACCACGTCCC TGGGACACGC CTGTCCCCC GGTGTACGCC GAGACCAAGC 1050
 ACTTCCCTTA CTCTCAGGC GACAAGGAGC AGCTGCGGCC CTCTTCTTA CTCAGCTCTC TGAGGCCACG 1120
 CCTGACTGGC GCTCGGAGGC TCGTGAGAC CATCTTCTG GGTTCAGGC CCTGGATGCC AGGGACTCCC 1190
 CGCAGGTTGC CCGGCTGCC CCAGCGCTAC TGGCAATGC GCGCCCTGTT TCTGGAGCTG CTTGGGAACC 1260
 ACGCGCATGG CCCCTACGGG GTGCTCTCA AGACGCACTG CCGGTCGGA GCTGCGGTCA CCCCAGCAGC 1330
 CGGTGTCTGT GCGCGGAGA AGCCCCAGGG CTCTGTGGCG GCGCGGAGG AGGAGGACAC AGACCCCGT 1400
 CGCTGTGTGC AGCTGCTCCG CCAGCAGAGC AGCCCCGGC AGGTGTACGG CTTGTGCGG GCCTGCCTGC 1470
 GCGGCTGGT GCGCCAGGC ATGCGGAGC CTCTGGGCT CCAGGCACAA CGAACGCGC TTCCTCAGGA ACACCAAGAA 1540
 GTTCATCTCC CTGGGGAAGC ATGCCAAGCT CTCGCTGAG GAGCTGACGT GGAAGATGAG CGTGCGGAC 1610
 TGGCGTTGGC TGGCGAGGAG CCCAGGGGTT GCGTGTGTT CCGCGCGAGA GCACCGTCTG CGTGAGGAGA 1680
 TCCTGGCCAA GTTCTGCAC TGGCTGATGA GTGTGTACGT CGTGAGCTG CTCAGGTCTT TCTTTATGT 1750
 CACGGAGACC ACGTTTCAA AGAACAGGCT CTTTTCTAC CGGAAGAGTG TCTGGAGCAA GTTGCAAGC 1820
 ATTGAATCA GACAGCACTT GAAGAGGGTG CAGCTGCGGG AGCTGTGCGA AGCAGAGTGC AGGCAGCATC 1890
 GGGAAAGCAG GCGCGCCCTG CTGACGTCCA GACTCCGCTT CATCCCCAAG CCTGACGGGC TGCGGCCGAT 1960
 TGTGAACATG GACTACGTCG TGGGAGCCAG AACGTTCCG CAGGACACCA TCCCTCAGAG TCTCAGCTG 2030
 AGGGTGAAG CACTGTTTCA CGTGCTCAAC TACGAGCGGG GCGCGCGCCC CCGGCTCCTG GCGGCTCTG 2100
 TGCTGGGCTT GAGCAGATATC CACAGGGGCT GCGCACCTT CGTGCTGCGT GTGCGGGCCC AGGACCCGCC 2170
 GCCTGAGCTG TACTTTGTCA AGGTGGATGT GACGAGCTC CTCACCCAG CGAAACCTT CCTCAGGACC CTGGTCCGAG 2240
 GAGGTATCG CCAGCATCAT CAAACCCAG AACACGTACT GCGTGCCTG GTATGCCGTG GTCCAGAAGG 2310
 CCGCCCATGG GCACGTCCG AAGGCCITCA AGAGCCACGT CTCTACCTTG ACAGACCTCC AGCCGTACAT 2380
 GCGACAGTTC GTGGCTCACC TGCAGGAGC CAGCCGCTG AGGGATGCGG TCGTCATCGA CAGAGCTCC 2450
 TCCTGAATG AGGCCAGCAG TGGCTCTTC GACGTCTTC TACGTTTAT GTGCCACCAC GCGGTGCGCA 2520
 TCAGGGGCAA GTCCTACGTC CAGTGCCAGG GGATCCCGCA GGGCTCCATC CTCTCCACGC TGTCTGTCAG 2590
 CCTGTGCTAC GCGCAGATGG AGAACAGCT GTTTGCGGGG ATTGCGCGGG ACGGGCTGCT CCGGCTTTG 2660
 GTGGATGATT TCTTGTGGT GACACCTCAC CTCACCCAG CGAAACCTT CCTCAGGACC CTGGTCCGAG 2730
 GTGTCCCTGA GTATGGCTGC GTGGTGAAT TCGGGAAGC AGTGGTGAAC TTCCCTGTAG AAGACGAGGC 2800
 CCTGGGTGGC ACGGCTTTTG TTCAGATGCC GCGCCACGGC CTATTCCCTT GGTGCGGCTT GCTGCTGGAT 2870
 ACCCGGACCC TGGAGGTGCA GAGCGACTAC TCCAGCTATG CCGGACCTC CATCAGAGCC AGTCTCACCT 2940
 TCAACCGCGG CTTCAAGGCT GGGAGGAACA TCGTTCGCAA ACTCTTTGGG GTCTTGCGGC TGAAGTGTCA 3010
 CAGCCTGTTT CTGGATTTC AGGTGAACAG CCTCCAGAG GTGTGCACCA ACATCTACAA GATCCTCTG 3080
 CTGAGGCGT ACAGGTTTCA CGCATGTGT CTGCACTCC CATTTCATCA GCAAGTTTGG AAGAACCCCA 3150
 CATTTTCTT GCGGCTCATC TCTGACACGG CCTCCCTCTG CTACTCCATC CTGAAAGCCA AGAACGAGG 3220
 GATGTGCTG GGGGCCAAGG GCGCGCGCGG CCTCTGCCC TCCGAGGCGG TGCAGTGGCT GTGCCACCA 3290
 GCATTCTGTC TCAAGCTGAC TCGACACCGT GTCACCTACG TGCCACTCCT GGGGTCACTC AGGACAGCCC 3360
 AGACGAGCT GAGTCGGAAG CTCCCCGGGA CGACGCTGAC TGCCCTGGAG GCGCGAGCCA ACCCGGCACT 3430
 GCCCTCAGAC TTCAAGACCA TCCTGGAAGT ATGGCCACCC GCGCCACAGC AGGCGGAGAG CAGACACAG 3500
 CAGCCCTGTC ACGCCGGGCT CTACGTCCCA GGGAGGGAGG GCGCGGCCAC ACCCAGGCC GCACCGCTGG 3570
 GAGTCTGAG CTTGAGTGAG TGTTTGGCG AGGCTGCGAT GTCCGGCTGA AGGCTGAGTG TCCGGCTGAG 3640
 GCCTGAGCGA GTGTCCAGCC AAGGGCTGAG TGTCCAGCAC ACCTGCCGTC TTCACTTCCC CACAGGCTGG 3710
 CGCTCGGCTC CACCCAGGG CCAGCTTTTC CTCACAGGA GCGCGGCTTC CACTCCCCAC ATAGGAATAG 3780
 TCCATCCCCA GATTCCCAT TGTTCACCCC TCGCCCTGCC CTCTTTGCC TTCCACCCCC ACCATCCAGG 3850
 TGGAGACCT GAGAAGGACC CTGGGAGCTC TGGGAATTG GAGTGACCAA AGGTGTGCC TGTACACAGG 3920
 CGAGGACCT GCACCTGGAT GGGGCTCCCT GTGGGTCAA TTGGGGGGAG GTGCTGTGG AGTAAATAC 3990
 TGAATATAG AGTTTTTCAG TTTTGA AAAA AAAAAA AA 4042

Fig. 7

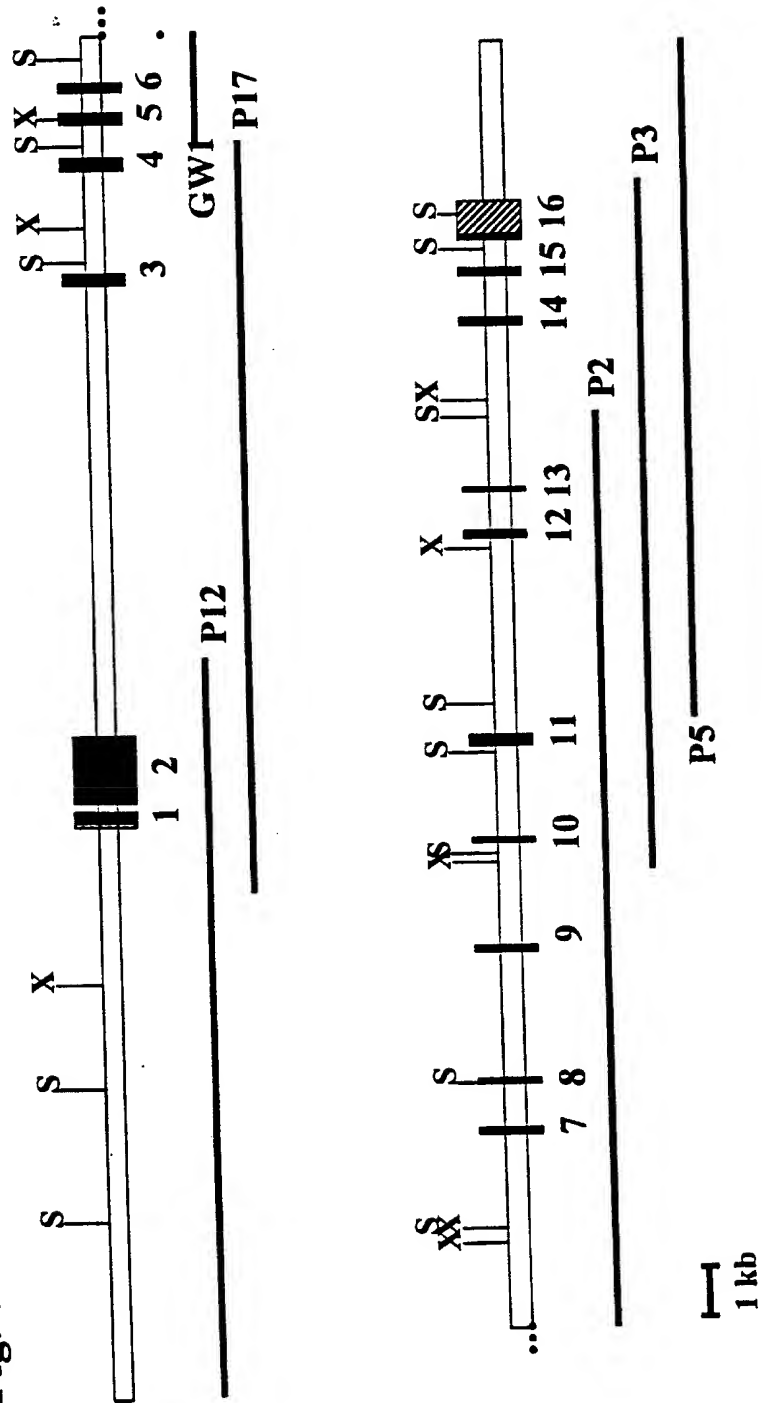


Fig. 8A

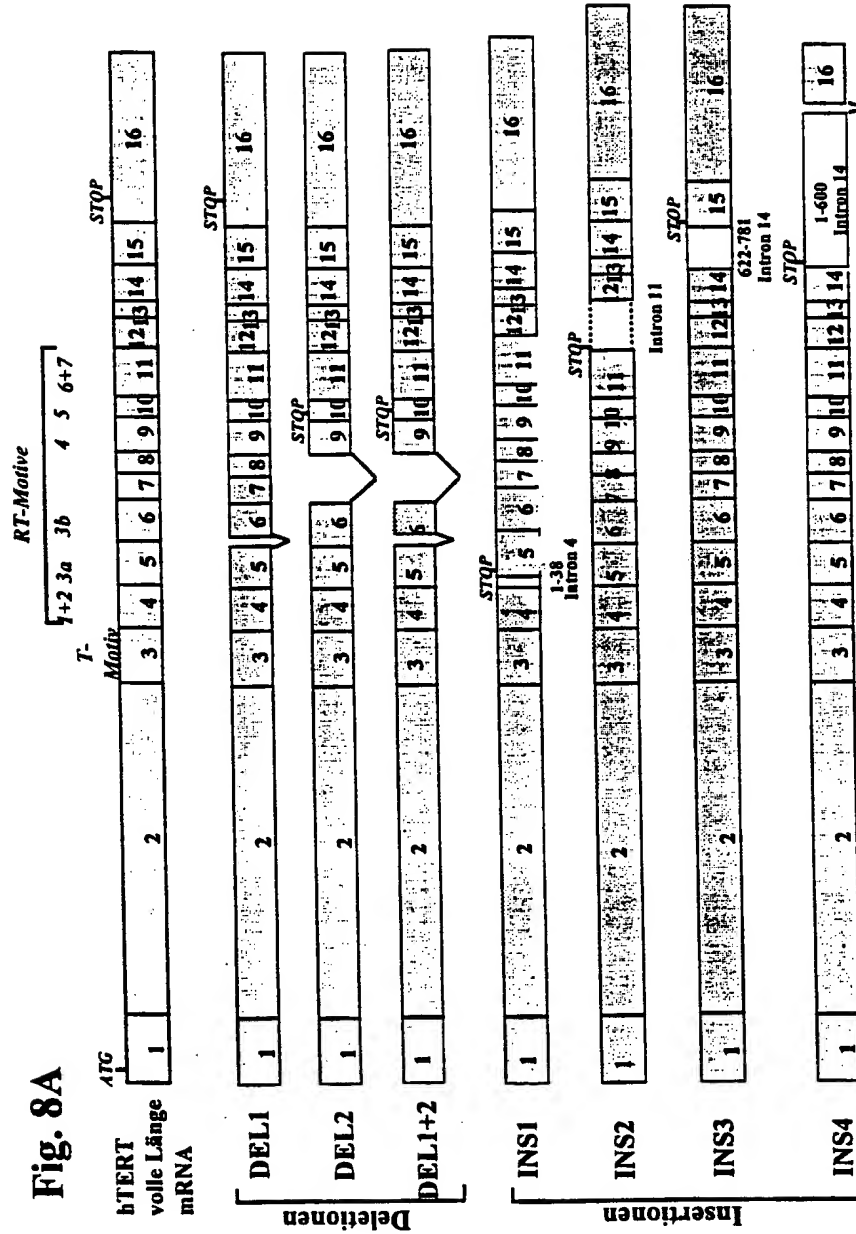
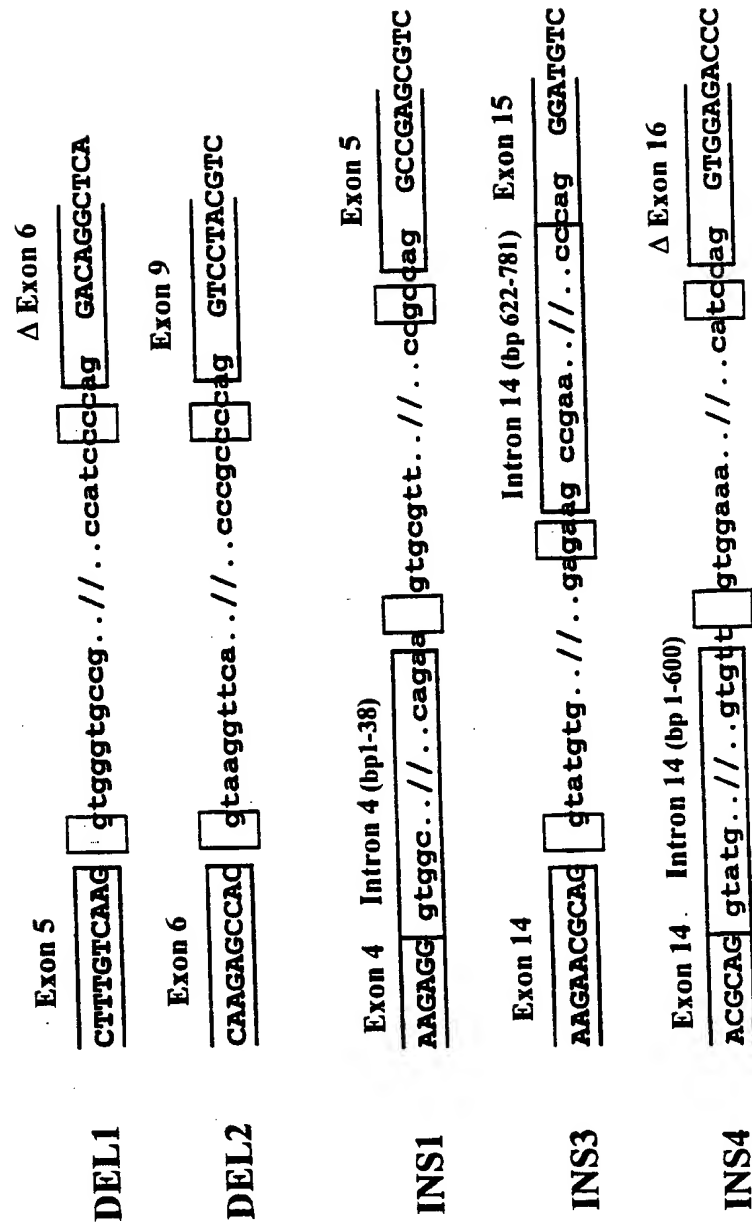


Fig. 8B



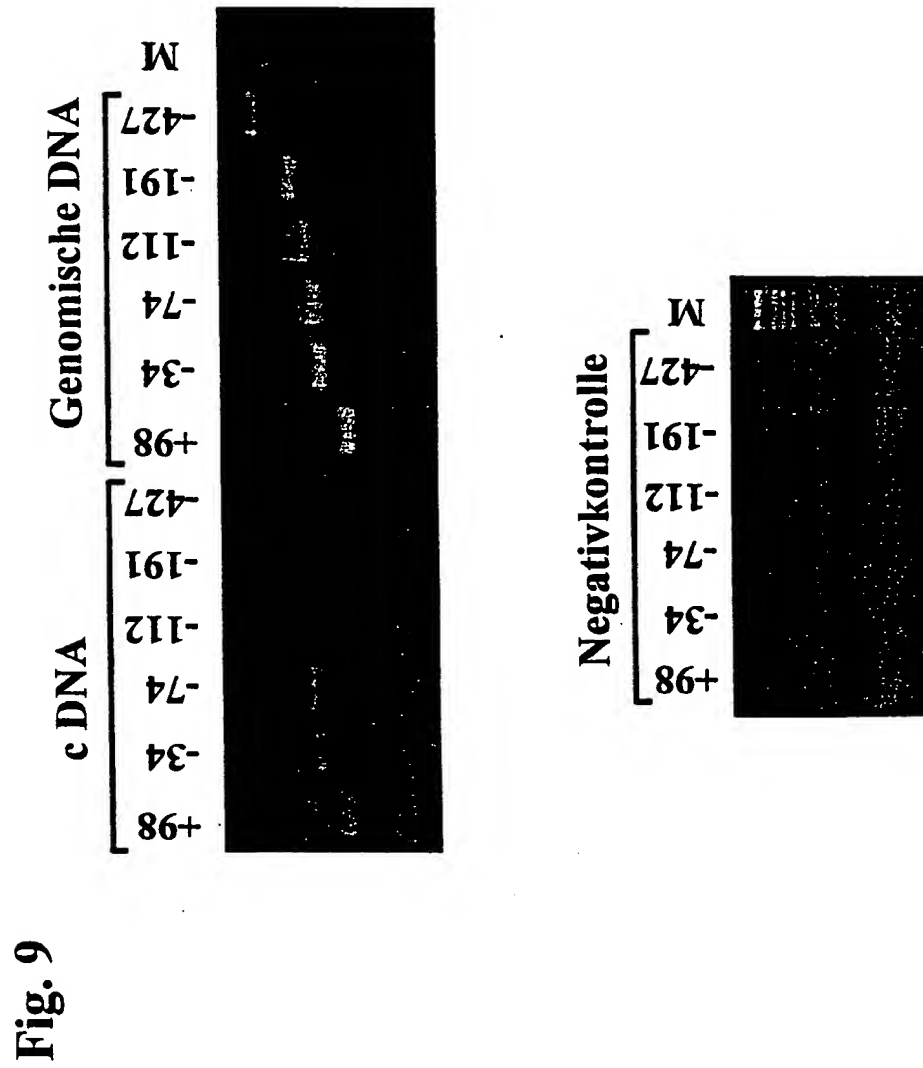


Fig. 10

ACTTGAGCCC AAGAGTTCAA GGCTACGGTG AGCCATGATT GCAACACCAC ACGCCAGCCT TGGTGACAGA -11204
 ATGAGACCCT GTCTCAAAAA AAAAAAAAAA AATTGAAATA ATATAAGCA TCTTCTCTGG CCACAGTGGA -11134
 ACAAAAACAG AATCAACAA CAAGAGGAAT TTTGAAAACT ATACAAACAC ATGAAAAATTA AACAAATATAC -11064
 TTCTGAATGA CCAGTGAGTC AATGAAGAAA TTAATAAGGA AATTGAAAAA TTTATTTAAG CAAATGATAA -10994
 CGGAAACATA ACCTCTCAAA ACCCAGCGTA TACAGCAAAA GCAGTGCTAA GAAGGAAGTT TATAGCTATA -10924
 AGCAGCTACA TCAAAAAAGT AGAAAAGCCA GGCGCAGTGG CTCATGCCTG TAATCCCAGC ACTTTGGGAG -10854
 GCCAAGGCGG GCAGATCGCC TGAGGTGAGG AGTTCGAGAC CAGCCTGACC AACACAGAGA AACCTTGTCG -10784
 CTACTAAAAA TACAAAAATTA GCTGGGCATG GTGGCACATG CCTGTAATCC CAGCTACTCG GGAGGCTGAG -10714
 GCAGGATAAC CGCTTGAAAC CAGGAGGTGG AGGTTGCGGT GAGCCGGGAT TGGCCCATTT GACTCCAGCC -10644
 TGGGTAACAA GAGTGAAACC CTGTCTCAAG AAAAAAAAAA AAGTAGAAAA ACTTAAAAAT ACAACCTAAT -10574
 GATGCACCTT AAAGAAGTAG AAAAGCAAGA GCAAACTAAA CCTAAAATTG GTAAAAGAAA AGAATAATA -10504
 AAGATCAGAG CAGAAATAAA TGAAACTGAA AGATAACAAT ACAAAGATC AACAAAAATTA AAAGTTGGTT -10434
 TTTTAAAAAG ATAAACAAAA TTGACAAACC TTGCCCAGA CTAAGAAAAA AGGAAAGAAAG ACCTAAATAA -10364
 ATAAAGTCAG AGATGAAAAA AGAGACATTA CAACTGATAC CACAGAAATT CAAAGGATCA CTAGAGGCTA -10294
 CTATGAGCAA CTGTACACTA ATAAATTGAA AAACCTAGAA AAAATAGATA AATTCTCTAGA TGCATACAAC -10224
 CTACCAAGAT TGAACCATGA AGAAATCCAA AGCCCAACAC GACCAATAAC AATAATGGGA TTAAGCCCAT -10154
 AATAAAAAAGT CTCTAGCAA AGAGAAGCCC AGGACCCAAT GGCTTCCCTG CTGGATTTTA CCAATCATTT -10084
 AAAGAAGAAAT GAATTCATCT CTTACTCAA CTATTCTGAA AAATAGAGGA AAGAATACTT CCAAACTCAT -10014
 TCTACATGGC CAGTATTACC CTGATTCCAA AACCAGACAA AACACATCA AAAACAAACA AACAAAAAAA -9944
 CAGAAAGAAA GAAACTACA GGCCAAATATC CCTGATGAAT ACTGATACAA AAATCCTCAA CAAAACACTA -9874
 GCAAAACAAA TTAACAACA CCTTCGAAAG ATCATTCTT GTGATCAAGT GGGATTTATT CCAGGGATGG -9804
 AAGGATGGTT CAACATATGC AAATCAATCA ATGTGATACA TCATCCCAAC AAAATGAAGT ACAAACCTA -9734
 TATGATTATT TCACCTTATG CAGAAAAAGC ATTTGATAAA ATTCTGCACC CTTCTATGATA AAAACCTCTA -9664
 AAAAAACGAG TATACAAGAA ACATACAGGC CAGGCACAGT GGCTCACACC TGGCATCCCA GCACCTCTGG -9594
 AGGCCAAGGT GGGATGATTG CTTGGGCCCC GAGGTTTGG ACTAGCCTGG GCAACAAAAA GAGACCTGGT -9524
 CTACAAAAAA CTTTTTTTAA AAATTAGCCA GGCATGATGG CATATGCCTG TAGTCCCAGC TAGTCTGGAG -9454
 GGTGAGGTGG GAGAATCACT TAAGCCTAGG AGGTCGAGGC TGCAGTGAGC CATGAACATG TCACTGTACT -9384
 CCAGCCTAGA CAACAGACA AGACCCCACT GAATAAGAA AGAGGAGAAG AGAAGGGAGA AGGGAGGGAG -9314
 AAGGGAGGAG GAGGAGAAGG AGGAGGTGGA GGAGAAGTGG AAGGGGAAGG GGAAGGGAAA GAGGAAGAAG -9244
 AAGAAACATA TTTCAACATA ATAAAGCCC TATATGACAG ACCGAGGTAG TATTATGAGG AAAAAGCTGA -9174
 AGCCTTTCTCT CTAAGATCTG GAAAATGACA AGGGCCCACT TTCACCACTG TGATTCAACA TAGTACTAGA -9104
 AGTCCTAGCT AGAGCAATCA GATAAGGAA AGAAATAAAA GGCATCCAAA CTGGAAAGGA AGAAGTCAAA -9034
 TTATCCTGTT TGCAGATGAT ATGATCTTAT ATCTGGAAAA GACTTAAGAC ACCACTAAAA AACTATTAGA -8964
 GGTGAATTTT GGTACAGCAG GATACAAAAT CAATGTACAA AAATCAGTAG TATTTCTATA TTCCAAACAGC -8894
 AAACAATCTG AAAAAGAAAC CAAAAAGCA GCTACAAATA AAATTAACA GCTAGGAATT AACCAAAGAA -8824
 GTGAAAGATC TCTACAATGA AAATATAAA ATGTTGATAA AAGAAATTTGA AGAGGGCACA AAAAAGAAAA -8754
 AGATATTCCA TGTTTATAGA TTGGAAGAAT AAATACTGTT AAAATGTCCA TACTACCCAA AGCAATTTAC -8684
 AAATTCATG CAATCCCTAT TAAATACTA ATGACGTTCT TCACAGAAAT AGAAGAAAAA ATTCTAAGAT -8614
 TTGTACAGAA CCACAAAAGA CCCAGAATAG CCAAAGCTAT CCTGACCAA AAGAACAAAA CTGGAAGCAT -8544
 CACATTACCT GACTTCAAAT TATACTACAA AGCTATAGTA ACCCAAATA CATGGTACTG GCATAAAAAAC -8474
 AGATGAGACA TGGACCAGAG GAACAGAATA GAGAATCCAG AAACAAATCC ATGCATCTAC AGTGAACCTA -8404
 TTTTTCACAA AGGTGCCAAG AACATACTTT GGGGAAAAAG TAATCTCTTC AATAAATGGT GCTGGAGGAA -8334
 CTGGATATCC ATATGCAAAA TAACAATATC AGAAGCTCTG CTCTCACCAT ATACAAAAAGC AAATCAAAAT -8264
 GGATGAAAGG CTTAAATCTA AAACCTCAAA CTTTGCAACT ACTAAAAGAA AACACCGGAG AAATCTCTCA -8194
 GGACATTGGA GTGGGCAAG ACTTCTTGAG TAATCCCTG CAGGCACAGG CAACCAAGC AAAACAGAC -8124
 AAATGGGATC ATATCAAGTT AAAAAGCTTC TGCCAGCAAA AGGAAACAAT CAACAAGAG CAAGAGACAAC -8054
 CCACAGAATG GGAGAATATA TTTGCAAACT ATTCACTTAA CAAGGAATTA ATAACCAGTA TATATAAGGA -7984
 GCTCAAACTA CTCTATAAGA AAAACACCTA ATAAGCTGAT TTTCAAAAAT AAGCAAAAGA TCTGGGTAGA -7914
 CATTCTCTCA AATAAGTCAT ACAAATGGCA AACAGGCATC TGAAAATGTG CTCAACACCA CTGATCATCA -7844
 GAGAAATGCA AATCAAACT ACTATGAGAG ATCATCTCAT CCCAGTTAAA ATGGCTTTTA TTCAAAGAC -7774
 AGGCAATAAC AAATGCCAGT GAGGATGTGG ATAAAAGGAA ACCCTTGGAC ACTGTTGGTG GGAATGGAAA -7704
 TTGCTACCAC TATGGAGAAC AGTTTGAAAG TTCCTCAAAA AACTAAAAAT AAAGCTACCA TACAGCAATC -7634
 CCATTGCTAG GTATATATCT CAAAAAAGGG AATCAGTGTA TCAACAAGCT ATCTCCACTC CCACATTAC -7564
 TGCAGCACTG TTTATAGCAG CCAAGGTTTG GAAGCAACCT CAGTGTCCAT CAACAGACGA ATGGAAAAAG -7494
 AAAATGTGGT GCACATACAC AATGGAGTAC TACGCAGCCA TAAAAAGAA TGAGATCCTG TCAGTTGCAA -7424
 CAGCATGGGG GGCACCTGGT AGTATGTTAA GTGAAATAAG CCAGGCACAG AAAGACAAAC TTTTCATGTT -7354
 CTCCCTTACT TGTGGGAGCA AAAATTAATA CAATTGACAT AGAATAGAG GAGAATGGTG GTTCTAGAGG -7284
 GGTGGGGGAC AGGGTGACTA GAGTCAACAA TAATTTTATTG TATGTTTTAA AATAACTAAA AGAGTATAAT -7214
 TGGGTTGTTT GTAACACAAA GAAAGGATAA ATGCTTGAAG GTGACAGATA CCCCATTTAC CTGATGTGA -7144
 TTATTACACA TTGTATGCCT GTATCAAAAT ATCTCATGTA TGCTATAGAT ATAAACCCTA CTATATTAAT -7074
 AATTAATAAT TTAATGGCCA GGCACGGTGG CTCATGTCCG TAATCCCAGC ACTTTGGGAG GCCGAGGCGG -7004
 GTGGATCACC TGAGGTCAGG AGTTTGAAC CAGCTGGCC ACCATGATGA AACCTGTCT CTACTAAAGA -6934
 TACAAAAAAT AGCCAGCGGT GGTGGCACAT ACCTGTAGTC CCAACTACTC AGGAGGCTGA GACAGGAGAA -6864
 TTGCTTGAAC CTGGGAGGCG GAGGTTGCAG TGAGCCGAGA TCATGCCACT GCACTGCAGC CTGGGTGACA -6794
 GAGCAAGACT CCATCTCAAA AAAAAACAA AAAAAAGAA ATTAAAAAG GTAATTAACC ACTTAATCTA -6654
 AATATATACT CTACTATATT AGAAGTTAAA AATTAAAAA ATTAAAAAG GTAATTAACC ACTTAATCTA -6654
 AAATAAGAAC AATGTATGTG GGGTTTCTAG CTTCTGAAGA AGTAAAAGTT ATGGCCACGA TGGCAGAAAT -6584

ERSATZBLATT (REGEL 26)

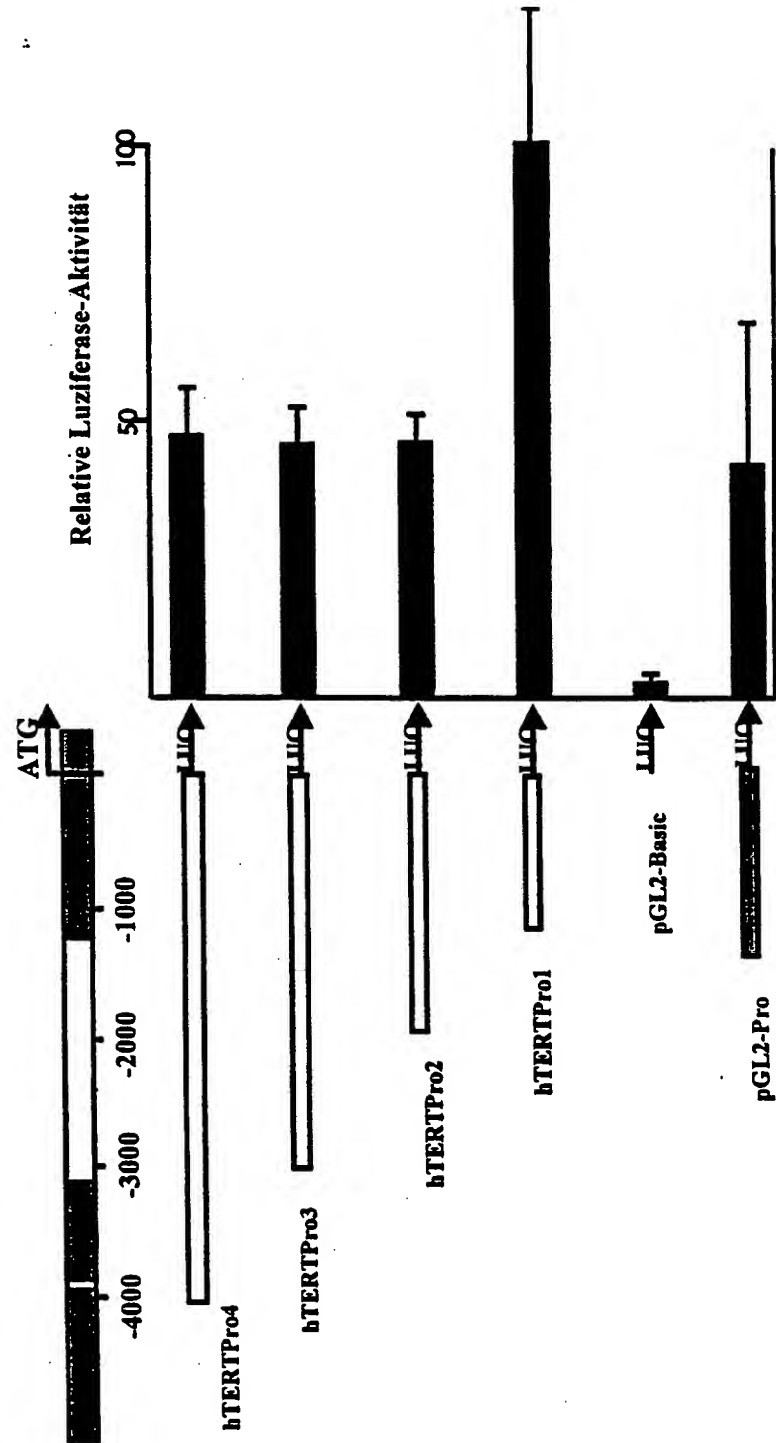
Fig. 10

GTGAGGAGGG AACAGTGGAA GTTACTGTTG TTAGACGCTC ATACTCTCTG TAAGTGACTT AATTTTAACC -6514
 AAAGACAGGC TGGGAGAAGT TAAAGAGGCA TTCTATAAGC CCTAAAACAA CTGCTAATAA TGGTGAAAGG -6444
 TAATCTCTAT TAATTACCAA TAATTACAGA TATCTCTAAA ATCGAGCTGC AGAATTGGCA CGTCTGATCA -6374
 CACCGTCTCT TCATTACAGG TGCTTTTTTT CTTGTGTGCT TGGAGATTTT CGATTGTGTG TTCGTGTTTG -6304
 GTTAAACTTA ATCTGTATGA ATCCTGAAAC GAAAAATGGT GGTGATTTCC TCCAGAAGAA TTAGAGTACC -6234
 TGGCAGGAAG CAGGTGGCTC TGTGGACCTG AGCCACTCA ATCTTCAAGG GTCTCTGGCC AAGACCCAGG -6164
 TGCAAGGCAG AGGCCTGATG ACCCGAGGAC AGGAAAGCTC GGATGGGAAG GGGCGATGAG AAGCCTGCCT -6094
 CGTTGGTGAG CAGCGCATGA AGTGCCCTTA TTTACGCTTT GCAAAGATTG CTCTGGATAC CATCTGGAAG -6024
 AGGCGGCCAG CGGGAATGCA AGGAGTCAGA AGCCTCCTGC TCAAACCCAG GCCAGCAGCT ATGGCGCCCA -5954
 CCCGGCGGTG TGCCAGAGGG AGAGGAGTCA AGGCACCTCG AAGTATGGCT TAAATCTTTT TTTACCTTGA -5884
 AGCAGTGACC AAGGTGTATT CTGAGGGAAG CTTGAGTTAG GTGCCCTTCT TAAACAGAA AGTCATGGAA -5814
 GCACCTTTCT CAAGGGA AAAACAGGAA CCAGACGCCC GCTCTGCGGT CATTTACCTC TTTCTCTCTT CCCTCTCTTG -5744
 CCCTCGCGGT TTCTGATCGG GACAGAGTGA CCCCGCTGGA GCTTCTCCGA GCCCGTGTG AGGACCCCTC -5674
 TGCAAGGGC TCCACAGACC CCCGCCCTGG AGAGAGGAGT CTGAGCCTGG CTTAATAACA AACTGGGATG -5604
 TGGCTGGGGG CGGACAGCGA CGGCGGGATT CAAAGACTTA ATTCATGAG TAAATCAAC CTTTCCACAT -5534
 CCGAATGGAT TTGGATTTTA TCTTAATATT TTCTAAATT TCATCAATA ACATTGAGGA GTGCAAAAT -5464
 CCAAAGCGT AAAACAGGAA CTGAGCTATG TTTGCCAAGG TCCAAAGGACT TAATAACCAT GTTCAGAGGG -5394
 ATTTTTCCGC CTAAGTACTT TTTATTGGTT TTCATAAGGT GGCTTAGGGT GCAAGGGAAA GTACACGAGG -5324
 AGAGGCTTGG GCGGCAGGGC TATGAGCAGC GCAGGGCCAC CGGGGAGAGA GTCCCCGGCC TGGGAGGCTG -5254
 ACAGCAGGAC CACTGACCGT CCTCCCTGGG AGCTGCCACA TTGGGCAACG CGAAGGCGGC CACGCTGCGT -5184
 GTGACTCAGG ACCCATACCC GGCTTCCCTG GCCCACCAC ACTAACCCAG GAAGTCACGG AGCTCTGAAC -5114
 CCGTGGAAAC GAACATGACC CTTGCCTGCC TGCTTCCCTG GGTGGGTCAA GGGTAATGAA GTGGTGTGCA -5044
 GGAATGGGCC ATGTAAATTA CACGACTCTG CTGATGGGGA CCGTTCTTTC CATCATTTAT CATCTTCCAC -4974
 CCCAAGGACT GAATGATTCC AGCAACTTCT TCGGGTGTGA CAAGCCATGA CAAAACCTCAG TACAAACACC -4904
 ACTCTTTTAC TAGGCCACCA GAGCAGGSCC CACACCCCTG ATATATTAAG AGTCCAGGAG AGATGAGGCT -4834
 GCTTTTCAGC ACCAGGCTGG GGTGACAACA GCGGCTGAAC AGTCTGTTC TCTAGACTAG TAGACCCCTG -4764
 CAGGCACTCC CCCAGATTCT AGGGCCTGGT TGCTGCTTCC CGAGGGCGCC ATCTGCCCTG GAGACTCAGC -4694
 CTGGGTGCC AACTGAGGC CAGCCCTGTC TCCACACCCT CCGCCTCCAG GCCTCAGCTT CTCCAGCAGC -4624
 TTCTAAACC CTGGGTGGC CGTGTCCAG CGCTACTGTC TCACCTGTCC CACTGTGTCT TGTCTCAGC -4554
 ACGTAGCTCG CACGGTTCCT CCTCACATGG GGTGTCTGTC TCCTTCCCCA ACACCTACAT CGCTTGAAGG -4484
 GAGGAGATTG TGCGCCTCCC AGACTGGCTC CTCTGAGCCT GAACCTGGCT CGTGGCCCCC GATGAGGTT -4414
 CCTGGCGTCC GGCTGCACGC TGACCTCCAT TTCCAGGCGC TCCCGTCTC CTGTCTCATG CCGGGGCTG -4344
 CCGGTGTGTT CTCTGTGTTT TGTGCTCCTT TCCAGCTCA GCTGCGTGTG TCTCTGCGCG CTAGGCTCTC -4274
 GGGGTTTTTA TAGGCATAGG ACGGGGCGT GGTGGGCCAG GCGCTCTTG GGAATGCAA CATTTGGGTG -4204
 TGAAAGTAGG AGTGCTGTG CTCACCTAGG TCCACGGGCA CAGGCTGGG GATGGAGCCC CCGCCAGGGA -4134
 CCCGCCCTTC TCTGCCCAGC ACTTTCCTGC CCCCTCCCT CTGGAACACA GAGTGGCAGT TTCCACAAGC -4064
 ACTAAGCATC CTCTTCCCAA AAGACCCAGC ATTGGCACCC CTGACATTT GCCCCACAGC CCTGGGAATT -3994

c-Myc
 CACGTGACTA CGCACATCAT GTACACACTC CCGTCCACGA CCGACCCCG CTGTTTTATT TTAATAGCTA -3924
 CAAAGCAGGG AAATCCCTGC TAAATGTCC TTTAACAAC TGGTTAAACA AACGGGTCCA TCCGCACGGT -3854
 GGACAGTTCC TCACAGTGAA GAGGAACATG CCGTTTATAA AGCCTGCAGG CATCTCAAGG GAATTACGCT -3784
 GAGTCAAAAC TGCCACCTCC ATGGGATACG TACGCAACAT GCTCAAAAAG AAAGAATTTC ACCCCATGGC -3714
 GGGGAGTGG TTAGGGGGGT TAAGGACGGT GGGGGCGGCA GCTGGGGGCT ACTGCACAGC CTTTTACTA -3644
 AAGCCAGTTT CTGTTTCTG ATGGTATTGG CTCAGTTATG GGAGACTAAC CATAGGGGAG TGGGGATGGG -3574
 GGAACCCGGA GGCTGTGCCA TCTTTGCCAT GCCCGAGTGT CTTGGGCAAG ATAATGCTCT AGAGATGCC -3504
 ACGTCTGAT TCCCCAAAC CTGTGGACAG AACCCTGCGG GCCCCAGGGC CTTTGCAGGT GTGATCTCCG -3434
 TGAGGACCT GAGGTCTGGG ATCCTTCGGG ACTACCTGCA GGCCCGAAAA GTAATCCAGG GGTCTGGGA -3364
 AGAGGCGGGC AGGAGGGTCA GAGGGGGGCA GCCTCAGGAC GATGGAGGCA GTCAGTCTGA GGCTGAAAG -3294
 GGAGGGAGGG CCTCGAGCCC AGGCCTGCAA GCGCTCCAG AAGCTGGAAA AAGCGGGGAA GGGACCTCC -3224
 ACGGAGCCTG CAGCAGGAAG GCACGGCTGG CCCTTAGCCC ACCAGGGCCC ATCGTGGACC TCCGGCTCC -3154
 GTGCCATAGG AGGGCACTCG CGCTGCCCTT CTAGCATGAA GTGTGTGGG ATTTGCAGAA GCAACAGGAA -3084
 ACCCATGCAC TGTGAATCTA GGATTATTTC AAAACAAGG TTTACAGAAA CATCCAAGGA CAGGGCTGAA -3014
 GTGCCTCCGG GCAAGGGCAG GGCAGGCAG AGTGATTTTA TTTAGCTATT TTATTTTATT TACTTACTTT -2944
 CTGAGACAGA GTTATGCTCT TGTGCCCCAG GCTGGAGTGC AGCGGCATGA TCTTGGCTCA CTGCAACCTC -2874
 CGTCTCCTGG GTTCAAGCAA TTCTCGTGCC TCAGCCTCCC AAGTAGCTGG GATTTCAGGC GTGCACCACC -2804
 ACACCCGGCT AATTTGTAT TTTTAGTAGA GATGGGCTTT CACCATGTTG GTCAGCTGA TCTCAAAATC -2734
 CTGACCTCAG GTGATCCGCC CACCTCAGCC TCCCAAAGTG CTGGGATTAC AGGCATGAGC CACTGCACCT -2664
 GGCCTATTTA ACCATTTTAA AACTTCCCTG GGCTCAAGTC ACACCCACTG GTAAGGAGTT CATGGAGTTC -2594
 AATTTCCCT TTAAGCAGGA GTTACCCTCC TTTGATATTT TCTGTAATTC TTCGTAGACT GGGGATACAC -2524
 CGTCTCTTGA CATATTCACA GTTCTGTGA CCACCTGTTA TCCCATGGGA CCCACTGCAG GGGCAGCTGG -2454
 GAGGCTGCAG GCTTCAGTTC CCAGTGGGGT TGCCATCTGC CAGTAGAAAC CTGATGTAGA ATCAGGGGCG -2384
 AAGTGTGGAC ACTGTCTGTA ATCTCAATGT CTCAGTGTGT GCTGAAACAT GTAGAAATTA AAGTCCATCC -2314
 CTCTACTCT ACTGGGATTG AGCCCCCTCC CTATCCCCC CCAGGGGCGAG AGGAGTTCCT CTCACTCCTG -2244
 TTGAGGAAGG AATGATACTT TGTATTTTT CACTGCTGGT ACTGAATCCA CTGTTTCATT TGTGTTGTTG -2174
 TTTGTTTTGT TTTGAGAGGC GGTTCCTCCT TTGTTGCTCA GGCTGGAGGG AGTGCAATG CCGCATCTTG -2104
 GCTTACTGCA GCCTCTGCCT CCCAGGTTCA AGTGATTCTC CTGCTTCCGC CTCCCATTTG GCTGGGATTA -2034
 CAGCACCCG CCACCATGCC CAGCTAATTT TTTGATTTT TAGTAGAGAC GGGGTGGGT GGGGTTTACC -1964

ATGTTGGCCA GGCTGGTCTC GAACCTCTGA CCTCAGATGA TCCACCTGCC TCTGCCTCCT AAAGTGCTGG -1894
GATTACAGGT GTGAGCCACC ATGCCCAGCT CAGAATTAC TCTGTTTGA AACATCTGGG TCTGAGGTAG -1824
GAAGCTCACC CCACTCAAGT GTTGTGGTGT TTTAACAAT-Box
GTTAAGCCAA TGATAGAATT TTTTATTGT TGTTAGAACA -1754
CTCTTGATGT TTTACACTGT GATGACTAAG ACATCATCAG CTTTTCAAAG ACACACTAAC TGCACCCATA -1684
ATACTGGGGT GTCTTCTGGG TATCAGCAAT CTTCATTGAA TGCCGGGAGG CGTTTCCTCG CCATGCACAT -1614
GGTGTAAAT ACTCCAGCAT AATCTTCTGC TTCCATTCT TCTCTCCCT CTTTAAAT TGTGTTTCT -1544
ATGTTGGCTT CTCTGCAGAG AACCAGTGTA AGCTACAACT TAACTTTTGT TGGAACAAAT TTTCCAACC -1474
Spi
GCCCTTTTG CCTAGTGGCA GAGACAATC ACAACACAG CCCTTTAAAG AGGCTTAGGG ATCACTAAGG -1404
GGATTCTAG AAGAGCGACC TGTAATCTTA AGTATTTACA AGACGAGGCT AACCTCCAGC GAGCGTGACA -1334
GCCCAGGGAG GGTGCGAGGC CTGTTCAAAT GCTAGCTCCA TAAATAAAGC AATTTCCTCC GGCAGTTTCT -1264
GAAAGTAGGA AAGGTTACAT TTAAGGTTGC GTTTGTTAGC ATTTCACTGT TTGCCGACCT CAGCTACAGC -1194
ATCCCTGCAA GGCCTCGGGA GACCCAGAAG TTTCTCGCCC CTTAGATCC AAACCTGAGC AACCCGGAGT -1124
CTGGATTCTT GGAAGTCTT CAGCTGTCCT GCGGTTGTGC CGGGGCCCA GGTCTGGAGG GGACCACTGG -1054
CCGTGTGGCT TCTACTGCTG GGCTGGAAAT CGGGCCTCCT AGCTCTGCAG TCCGAGGCTT GGAGCCAGGT -984
GCCTGGACCC CGAGGCTGCC CTCACCCCTG TCGGGGCGGG ATGTGACCAG ATGTTGGCCT CATCTGCCAG -914
ACAGAGTGCC GGGGCCCAGG GTCAAGGCCG TTGTGGCTGG TGTGAGGCGC CGGTGCGCG GCCAGCAGGA -844
GCGCCTGGCT CCATTCCCA CCCTTCTCG ACGGGAGCGC CCCGGTGGGT GATTAACAGA TTTGGGGTGG -774
TTTGCTCATG GTGGGACCC CTCGCCGCT GAGAACCTGC AAAGAGAAAT GACGGGCTTG TGTCAGGAG -704
CCCAAGTCGC GGGGAAGTGT TGCAGGGAGG CACTCCGGGA GGTCCCGCT GCGCGTCCAG GGAGCAATGC -634
GTCTCGGGT TCCTCCGAG CCGCTCTAC GCGCCTCCGT CCTCCCTTC ACCTCCGCA TTCGTGGTGC -564
CCGGAGCCCG ACGCCCGCG TCCGGACCTG GAGGCAGCCC TGGGTCTCG GATCAGGCCA GCGGCCAAAG -494
GGTCGCCGA CGCACTGTT CCCAGGCCT CCACATCATG GCCCCTCCT CGGGTTACCC CACAGCCTAG -424
GCCGATTCTA CCTCTCTCCG CTGGGGCCCT CGCTGGCGTC CTGCACCTT GGGAGCGCGA GCGGGCGG -354
GCGGGGGAAG CGCGGCCAG ACCCCCGGT CGGCCGGAG CAGCTGCGCT GTCGGGGCCA GGCCGGGCTC -284
CCAGTGGATT CGCGGGCACA GACGCCAGG ACCGCGCTCC CACGTGCG GAGGGAAGTGG GGACCCGGGC -214
ACCGTCTCTG CCCCTTACC TTCCAGTCC GCCTCCTCG CGCGGAGCCC GCCCGTCCC GACCCCTCCC -144
GGGTCCCCG CCCAGCCCC TCCGGGCCCT CCCAGCCCC CCCCTTCTT TCCGGGGCC GCCCCCTCC -74
TCGCGGCGCG AGTTTCAGGC AGCGCTGCGT CCTGTGCGC ACGTGGGAAG CCCTGGCCCC GGCCACCCCC -4
GCGATG

Fig.: 11



SEQUENZPROTOKOLL

<110> Bayer AG

5 <120> Regulatorische DNA-Sequenzen aus der Si-Region vom Gen
der humanen katalytischen Telomerase-Untereinheit und
deren diagnostische und therapeutische Verwendung

10 <130> LeA32805-Ausland

<140>

<141>

15 <160> 20

<170> PatentIn Vers. 2.0

<210> 1

20 <211> 5126

<212> DNA

<213> Homo sapiens

<400> 1

25 gagctctgaa ccgtggaaac gaacatgacc cttgcctgcc tgcttccctg ggtgggtcaa 60
gggtaatgaa gtgggtgtgca ggaaatggcc atgtaaatta cacgactctg ctgatgggga 120
ccgttccttc catcattatt catcttcacc cccaaggact gaatgattcc agcaacttct 180
tcgggtgtga caagccatga caaaactcag tacaacacacc actcttttac taggcccaca 240
gagcacgggc cacaccctg atatattaag agtccaggag agatgaggct gctttcagcc 300
accaggctgg ggtgacaaca gcggctgaac agtctgttcc tctagactag tagaccctgg 360
30 caggcactcc cccaaattct agggcctggg tgctgcttcc cgaggggcgc atctgcctcg 420
gagactcagc ctgggggtgcc acactgaggc cagccctgtc tccacaccct ccgcctccag 480
gcctcagctt ctccagcagc ttcctaaacc ctgggtgggc cgtgttccag cgctactgtc 540
tcacctgtcc cactgtgtct tgtctcagcg acgtagctcg cacggttcct cctcacatgg 600
ggtgtctgtc tcttcccca acactcacat gcgttgaagg gaggagattc tgcgcctccc 660
35 agactggctc ctctgagcct gaacctggct cgtggccccc gatgcagggt cctggcgctc 720
ggctgcacgc tgacctccat ttccaggcgc tcccgcgtctc ctgtcatctg ccggggcctg 780
ccgggtgtgt ctctctgttc tgtgctcctt tccacgtcca gctgcgtgtg tctctgccc 840
ctagggtctc ggggttttta taggcatagg acggggggcgt ggtggggcag ggcgctcttg 900
40 ggaaatgcaa catttgggtg tgaaagttag agtgccctgtc ctacactagg tccacgggca 960
caggcctggg gatggagccc ccgccaggga cccgcccttc tctgccagc acttctctgc 1020
ccccctccct ctggaacaca gagtggcagt ttccacaagc actaagcacc ctcttcccaa 1080
aagaccagc attggcaccc ctggacattt gcccacagc cctgggaatt cacgtgacta 1140
cgcacatcat gtacacactc ccgtccacga ccgacccccc ctgttttatt ttaatagcta 1200
caaagcaggg aaatccctgc taaatgtcc ttaacaaac tggttaaaca aacgggtcca 1260
45 tccgcacggg ggacagtccc tcacagtga gaggaacatg ccgtttataa agcctgcagg 1320
catctcaagg gaattacgct gagtcaaaac tgcacacctc atgggatacg tacgcaacat 1380
gctcaaaaag aaagaatttc accccatggc aggggagtggt ttaggggggt taaggacggg 1440
ggggggcgga gctgggggct actgcacgca ccttttacta aagccagttt cctgggtctg 1500
atgggtattg ctcatgtatg ggagactaac cataggggag tggggatggg ggaacccgga 1560
50 ggctgtgcca tctttgccat gcccgagtgt cctgggcagg ataagtctct agagatgcc 1620
acgtcctgat tcccccaaac ctgtggacag aaccgcggcg gcccaggggc ctttgagggt 1680
gtgatctccg tgaggaccct gaggtctggg atccttcggg actacctgca ggcccgaaaa 1740
gtaatccagg ggttctggga agaggcgggc aggggggtca gaggggggca gcctcaggac 1800
gatggaggca gtcagtctga ggctgaaaag ggaggggagg cctcgagccc aggcctgcaa 1860
55 gcgcctccag aagctggaaa aagcggggaa gggaccctcc accgagcctg cagcaggaa 1920
gcacggctgg cccttagccc accaggggcc atcgtggacc tccggcctcc gtgccatagg 1980
agggcactcg cgtgcccctt ctgacatgaa gtgtgtgggg atttgcagaa gcaacaggaa 2040
acccatgcac tgtgaatcta ggattatttc aaaacaaagg tttacagaaa catccaagga 2100
cagggtgtaa gtgcctccgg gcaagggcag ggcaggcacg agtgatttta ttagctatt 2160
60 ttattttatt tacttacttt ctgagacaga gttatgctct tgggtcccag gctggagtgc 2220
agcggcatga tcttggctca ctgcaacctc cgtctcctgg gttcaagcaa ttctcgtgcc 2280
tcagcctccc aagtagctgg gatttcaggc gtgcaccacc acaccggcct aattttgtat 2340
ttttagtaga gatggggctt caccatgttg gtcaaactga tctcaaaatc ctgacctcag 2400
gtgatccgcc cacctcagcc tcccaaagtg ctgggattac aggcattgag cactgcacct 2460

2 / 18

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

```

ggcctattta accattttta aacttccctg ggctcaagtc acaccactg gtaaggagtt 2520
catggagttc aatttccctt ttaactcagga gttaccctcc ttgtatattt tctgtaattc 2580
ttcgtagact ggggatacac cgtctcttga catattcaca gtttctgtga ccacctgtta 2640
tcccatggga cccactgcag gggcagctgg gaggtgcag gcttcaggte ccagtggggt 2700
tgccatctgc cagtagaaac ctgatgtaga atcagggcgc aagtgtggac actgtcctga 2760
atctcaatgt ctcaagtgtg gctgaaacat gtagaaatta aagtcacatcc ctccactact 2820
actgggattg agccccctcc ctatcccccc ccaggggcag agggagtctct ctcaactcctg 2880
tggagggaagg aatgatactt tgttattttt cactgctggt actgaatcca ctgtttcatt 2940
ttgtgggttg tttgttttgt tttgagaggc ggtttcactc ttgttgctca ggctggaggg 3000
agtcgaatgg cgcgactctg gcttactgca gcctctgcct ccagagttca agtgattctc 3060
ctgcttccgc ctcccatttg gctgggatta caggcaccgc ccaccatgcc cagctaattt 3120
ttgtattttt tagtagagac gggggtgggt ggggttcacc atgttggcca ggctggtctc 3180
gaacttctga cctcagatga tccacctgcc tctgcctcct aaagtgtcgg gattacaggt 3240
gtgagccacc atgcccagct cagaatttac tctgtttaga aacatctggg tctgaggtag 3300
gaagctcacc ccactcaagt gttgtggtgt ttaagccaa tgatagaatt tttttattgt 3360
tgttagaaca ctcttgatgt tttacactgt gatgactaag acatcatcag cttttcaaag 3420
acacactaac tgcacccata atactggggt gtctctcggg tatcagcaat ctccattgaa 3480
tgccgggagg cgtttcctcg ccactgcacat ggtgttaatt actccagcat aatcttctgc 3540
ttccatttct tctcttccct ctcttctcct tgtgttttct atgttggcct ctctgcagag 3600
aaccagtgtg agctacaact taacttttgt tggaaacaaat tttccaaacc gcccccttgc 3660
cctagtggca gagacaattc acaaacacag ccccttaaaa aggccttaggg atcactaagg 3720
ggatttctag aagagcgacc tgtaactcta agtatttaca agacgaggct aacctccagc 3780
gagcgtgaca gcccaggagc ggtgcgagcg ctgttcaaat gctagctcca taaataaagc 3840
aatttctctc gccagtttct gaaagtagga aaggttatat ttaaggttgc gtttgttagc 3900
atctcagttg ttgccgacct cagctacagc atccctgcaa ggcctcggga gaccagaag 3960
ttctctcgcc ccttagatcc aaacttgagc aaccggaggt ctggattcct ggggaagtct 4020
cagctgtcct gcggttgtgc cggggcccca ggtctggagg ggaccagtgg ccgtgtggct 4080
tctactgctg ggtcggaggt cgggctcctc agctctgcag tccgaggctt ggagccaggt 4140
gcctggaccc cgaggctgcc ctccaccctg tgcgggcggg atgtgaccag atgttggcct 4200
catctcgccg acagagtgcc ggggcccagg gtcaaggcgg ttgtggctgg tgtgaggcgc 4260
ccggtgcgcg gccagcagga gcgcctggct ccatttccca ccctttctcg acgggaccgc 4320
cccgtgggtt gattaacaga tttgggggtg tttgtctatg gtggggaccc ctgcgcgct 4380
gagaacctgc aaagagaaat gacgggcctg tgtcaaggag cccaagtcgc ggggaagtgt 4440
tgcaggaggc caactccggga ggtcccgcgt gcccgctcag ggagcaatgc gtccctcgggt 4500
tcgtccccag ccgcgtctac gcgcctcctc cctccccctc acgtccggca ttcgtggtgc 4560
ccggagcccg acgccccgcg tccggacctg gaggcagccc tgggtctcgg gatcaggcca 4620
gcgcccaaag ggtcgccgca cgcacctgtt cccagggcct ccacatcatg gccccctcct 4680
cggttaccct cacagcctag gccgatcga cctctctcgg ctggggccct cgctggcgct 4740
cctgcacctt gggagcgca gcggcgcgcg ggcgggggag cgcggccag acccccggtt 4800
ccgcccggag cagctgcgct gtccggggcca ggccgggctc ccagtggatt cgcggggaca 4860
gacgcccagg accgcgctcc ccacgtggcg gagggactgg ggaccggggc acccgctctg 4920
ccccctcacc ttccagctcc gcctcctcgg cgcggacccc gccccgctcc gaccctccc 4980
gggtcccccg cccagccccc tccgggcccct cccagcccct ccccttctct tccgcggccc 5040
cgccctctcc tcgcggcgag agttttaggc agcgctgcgt cctgctgcgc acgtgggag 5100
ccctggcccc ggccaccccc gcgatg 5126
  
```

<210> 2
 <211> 4042
 <212> DNA
 <213> Homo sapiens

<400> 2

```

gtttcaggca gcgctgcgtc ctgctgcgca cgtgggaagc cctggccccg gccacccccg 60
cgatgccgcy cgctccccgc tgcgagcgcg tgcgtccctt gctgcgcagc cactaccgcy 120
aggtgctgcc gctggccacg ttcgtgcggc gcctggggcc ccagggtcgg cggctgggtc 180
agcgcgggga cccggcggtt ttccgcgcgc tgggtggcca gtgcctggtg tgcgtgcctc 240
gggacgcacg gccgcccccc gccgccccct ccttccgcca ggtgtcctgc ctgaaggagc 300
tgggtggccc agtgctgcag aggtgtgtcg agcgcggcgc gaagaacgtg ctggccttcg 360
gcttcgcgct gctggacggg gcccgcgggg gccccccca ggccctcacc accagcgtgc 420
gcagctacct gcccacacg gtgaccgacg cactgcgggg gagcggggcg tgggggctgc 480
tgctgcgccc cgtgggcgac gacgtgctgg ttacctgct ggacgcgtgc gcgctcttgc 540
tgctggtggc tcccagctgc gcctaccagg tgtgcgggccc gccgctgtac cagctcggcg 600
ctgccactca gggccggccc ccgccacacg ctagtggacc ccgaaggcgt ctgggatgcg 660
aacgggacct gaaccatagc gtcaggagag ccggggtccc cctgggacct ccagccccgg 720
gtgcgaggag gcgcgggggc agtgccagcc gaagtctgcc gttgcccaag agggccaggc 780
  
```


3 / 18

gtggcgctgc ccctgagccg gagcggagcg ccgttgggca ggggtcctgg gccaccccg 840
gcaggacgcg tggaccgagt gaccgtgggt tctgtgtggt gtcacctgcc agaccgccc 900
aagaagccac ctctttggag ggtgcgctct ctggcagcg cactccccc ccatccgtgg 960
gccgccagca ccacgcgggc ccccatcca catcgcgcc accacgtccc tgggacacgc 1020
5 cttgtccccc ggtgtacgcc gagaccaagc acttctctta ctctcaggc gacaaggagc 1080
agctgcggcc ctcccttcta ctacgtctc tgaggccag cctgactggc gctcggaggc 1140
tcgtggagac catctttctg ggttccaggc cctggatgcc agggactccc cgcaggttgc 1200
ccgcctgcc ccagcgctac tggcaaatgc ggcctctgt tctggagctg cttgggaacc 1260
acgcgcagtg cccctacggg gtgctcctca agacgcactg cccgtcgca gctgcggtca 1320
10 cccagcagc cggtgtctgt gcccgggaga agccccaggg ctctgtggcg gccccaggg 1380
aggaggacac agacccccgt cgctgtgtgc agctgtccc ccagcacagc agccccggc 1440
agggtgtacg cttcgtgccc gcctgctgct gcccgtggt gccccaggc ctctggggct 1500
ccaggcacaa cgaacgcgcg tctctcagga acaccaagaa gttcatctcc ctggggaagc 1560
atgccaaagt ctgcgtgacg gagctgacgt ggaagatgag cgtgcgggac tgcgtcgcc 1620
15 tgcgcaggag cccaggggtt ggctgtgttc cggccgcaga gcacctctg cgtgaggaga 1680
tcctggccaa gttcctgacg tggctgatga gtgtgtacgt cgtcagctg ctgaggctct 1740
ctttttatgt cagcgagacc acgtttcaaa agaacaggct ctttttctac cggaaagagt 1800
tctggagcaa gttgcaaaag attggaatca gacagcactt gaagagggtg cagctcgccg 1860
agctgtcggg agcagaggtc aggcagcatc gggaagccag gccgcctctg ctgacttcca 1920
20 gactccgctt catccccaag cctgacgggc tgcggccgat tgtgaacatg gactacgtcg 1980
tgggagccag aacgttccgc agagaaaaa gggccgagcg tctcacctcg aggtggaag 2040
cactgttccg cgtgctcaac tacgagcggg cgcggcgccc cggcctcctg ggcgctctg 2100
tgctgggcct ggacgatata cacaggccct ggccgacctt cgtgtgctgt gtgcccggcc 2160
aggacccgccc gcctgagctg tactttgtca aggtgagatg gacgggcccg tacgacacca 2220
25 tccccaggga caggctcacg gaggtcatcg ccagcatcat caaacccag aacacgtact 2280
gcgtgcgtcg gtatgccgtg gtccagaagg cgcgccatgg gcacgtcccg aaggccttca 2340
agagccacgt ctctaccttg acagacctcc agcctgatat gcgacagtcc gtggtcacc 2400
tgcaggagac cagcccgctg agggatgccc tctcatcga gcagagctcc tccctgaatg 2460
aggccagcag tggcctcttc gacgtcttcc tacgttctat gtgccaccac gccgtgcca 2520
30 tcaggggcaa gtccctacgtc cagtcccagg ggatcccga gggctccatc ctctccacgc 2580
tgctctgacg cctgtgtacg ggcgacatgg agaacaagct gtttgcgggg attcggcggg 2640
acgggctgct cctgcgtttg gtggatgatt tcttgttggg gacacctcac ctccaccagc 2700
cgaaaaacct cctcaggacc ctggtccgag gtgtccctga gtatggctgc gtggtgaact 2760
tgccgaagac agtgggtgaac ttccctgtag aagacgaggg cctgggtggc acggctttg 2820
35 ttccagatgc gcccacggc ctattccctt ggtgcggcct gctgctggat acccgaccc 2880
tggagggtgca gagcgactac tccagctatg cccggacctc catcagagcc agtctcacc 2940
tcaaccgccc ctccaaggct gggagggaaca tgcgtcgcaa actctttggg gtcttgcggc 3000
tgaagtgtca cagcctgttt ctggatttgc aggtgaacag cctccagacg gtgtgcacca 3060
40 acatctacaa gatcctcctg ctgcaggcgt acagggttca cgcattgtgt ctgcagctcc 3120
catttcatca cgaagtgttg aagaacccca catttttctt gcgcgtcatc tctgacacgg 3180
cctccctctg ctactccatc ctgaaagcca agaaccgagc gatgtcgctg ggggccaagg 3240
gcgcccggg cctctgccc tccgagggcg tgcagtggct gtgccacca gcatctctgc 3300
tcaagctgac tcgacaccgt gtccactacg tgccactcct ggggtcactc aggacagccc 3360
agacgcagct gagtgcggaag ctcccgggga cgcagctgac tgccctggag gccgcagcca 3420
45 accgggcact gcctcagac ttcaagacca tccgtgactg atggccaccc gccacagcc 3480
aggccgagag cagacaccag cagccctgtc acgcggggct ctacgtccca gggaggagg 3540
ggcggccccc acccaggccc gcaccgctgg gagtctgagg cctgagttag tgtttggccc 3600
aggcctgcat gtccggctga aggctgagtg tccggctgag gcctgagcga gtgtccagcc 3660
50 aagggctgag tgtccagcac acctgccgtc ttacttccc cacaggctgg cgctcggtc 3720
caccaccagg ccagcttttc ctaccaggga gcccggcttc cactccccac atagggaatag 3780
tccatcccca gatttcgcat tgttaccctc tcgcccctgc ctcttttgc ttccaccccc 3840
accatccagg tggagaccct gagaaggacc ctgggagctc tgggaatttg gagtgaacca 3900
agggtgtccc tgtacacagg caggagacct gcacctggat gggggctccc gtgggtcaaa 3960
55 ttggggggag gtgctgtggg agtaaaatac tgaatatatg agtttttcag ttttgaaaaa 4020
aaaaaaaaa aaaaaaaaaa aa 4042

<210> 3

<211> 11276

<212> DNA

<213> Homo sapiens

<400> 3

acttgagccc aagagttaa ggtacgggt agccatgatt gcaacaccac acgccagcct 60

tggtgacaga atgagaccct gtctcaaaaa aaaaaaaaaa aattgaaata atataaagca 120

tcttctcttg ccacagtga acaaaaccag aaatcaacaa caagaggaaat tttgaaact 180

	atataaacac	atgaaaatta	aacaatatat	ttctgaatga	ccagtgaagtc	aatgaagaaa	240
	ttaaaaagga	aattgaaaaa	tttatttaag	caaatgataa	cggaacata	acctctcaaa	300
	acccacggtta	tacagcaaaa	gcagtgtctaa	gaagggaagt	tatagctata	agcagctaca	360
	tcaaaaaagt	agaaaaagcca	ggcgagctgg	ctcatgcctg	taatcccagc	actttgggag	420
5	gccaaggcgg	gcagatcgcc	tgaggtcagg	agttcgagac	cagcctgacc	aacacagaga	480
	aaccttgtcg	ctactaaaaa	tacaaaatta	gctgggcatg	gtggcacatg	cctgtaatcc	540
	cagctactcg	ggaggtctgag	gcaggataac	cgcttgaacc	caggaggtgg	agggttcggt	600
	gagccgggat	tgcgccattg	gactccagcc	tgggtaacaa	gagtgaacc	ctgtctcaag	660
	aaaaaaaaaa	aagtagaaaa	acttaaaaaa	acaacctaat	gatgcacctt	aaagaactag	720
10	aaaagcaaga	gcaaaactaaa	cctaaaattg	gtaaaagaaa	agaaataata	aagatcagag	780
	cagaaataaaa	tgaaaactgaa	agataacaat	acaaaagatc	aacaaaatta	aaagtgtggt	840
	ttttgaaaag	ataaacaaaa	ttgacaaacc	tttgcccaga	ctaagaaaaa	aggaaaagaa	900
	acctaataaa	ataaagtcag	agatgaaaaa	agagacatta	caactgatac	cacagaat	960
	caaaggatca	ctagaggcta	ctatgagcaa	ctgtactata	ataaattgaa	aaacctagaa	1020
15	aaaatagata	aattcctaga	tgcatacaac	ctaccaagat	tgaacctatg	agaaatccaa	1080
	agcccaaaaa	gaccaataac	aataatggga	ctaaagccat	aataaaaaagt	ctcctagcaa	1140
	agagaagccc	aggacccaat	ggcttccctg	ctggatttta	ccaatcattt	aaagaagaat	1200
	gaattccaat	cctactcaaa	ctattctgaa	aaatagagga	aagaataact	ccaaactcat	1260
	tctacatggc	cagtattacc	ctgattccaa	aaccagacaa	aaacacatca	aaaacaaaca	1320
20	aacaaaaaaa	cagaaagaaa	gaaaactaca	ggccaatatc	cctgatgaat	actgatacaa	1380
	aaatcctcaa	caaaacacta	gcaaaacaaa	ttaaacaaca	ccttcgaaag	atcattcatt	1440
	gtgatcaagt	gggatttatt	ccagggatgg	aaggatggtt	caacatatgc	aatcaaatca	1500
	atgtgatata	tcatcccaac	aaaatgaagt	acaaaaacta	tatgattatt	tcactttatg	1560
	cagaaaaagc	atttgataaa	attctgcacc	tttcatgata	aaaaccctca	aaaaaccagg	1620
25	tatacaagaa	acatacaggc	caggcacagt	ggctcacacc	tgcgatccca	gcactctggg	1680
	aggccaaggt	gggatgattg	cttgggcccc	ggagtgttag	actagcctgg	gcaacaaaat	1740
	gagacctggt	ctacaaaaaa	ctttttttaa	aaattagcca	ggcatgatgg	catatgcctg	1800
	tagtcccagc	tagtctggag	gctgaggtgg	gagaatcact	taagcctagg	aggctcagggc	1860
	tgcagtgaag	catgaacatg	tcactgtact	ccagcctaga	caacagaaca	agaccccatc	1920
30	gaataagaag	aaggagaagg	agaagggaag	agggagggaag	aaggagggaag	gaggagaagg	1980
	aggaggtgga	ggagaagtg	aaggggaagg	ggaagggaag	gaggagaag	aagaaacata	2040
	tttcaacata	ataaaaagccc	tatatgacag	accgaggtag	tattatgagg	aaaaactgaa	2100
	agcctttcct	ctaagatctg	gaaaatgaca	agggcccact	ttcaccactg	tgattcaaca	2160
	tagtactaga	agtccctagc	agagcaatca	gataagagaa	agaaataaaa	ggcatccaaa	2220
35	ctgggaagga	agaaatcaaa	ttatcctgtt	tgcagatgat	atgatcttat	atctggaaaa	2280
	gacttaagac	accactaaaa	aactattaga	gctgaaattt	ggtacagcag	gatacaaaa	2340
	caatgtacaa	aaatcagtag	tattttctata	ttccaacagc	aaacaatctg	aaaaagaaac	2400
	caaaaaagca	gctacaaata	aaattaaaca	gctaggaatt	aaacaaagaa	gtgaaagatc	2460
40	tctacaatga	aaactataaa	atggttgata	aaagaaattg	agaggggcaca	aaaaaagaaa	2520
	agatattcca	tgttcataga	ttggaagaat	aaatactgtt	aaaatgtcca	tactacccaa	2580
	agcaattttac	aaattcaatg	caatccctat	taaaatacta	atgacgttct	tcacagaaat	2640
	agaagaaaca	attctaaagt	ttgtacagaa	ccacaaaaga	cccagaatag	ccaaagctat	2700
	cctgacccaa	aagaacaaaa	ctggaagcat	cacattacct	gacttcaaat	tatactacaa	2760
	agctatagta	acccaaacta	catggtactg	gcataaaaac	agatgagaca	tggaccagag	2820
45	gaacagaata	gagaatccag	aaacaaatcc	atgcatctac	agtgaactca	tttttgacaa	2880
	aggtgccaag	aacatacttt	ggggaaaaga	taatctcttc	aataaatggt	gctggaggaa	2940
	ctggatatcc	atatgcaaaa	taacaatact	agaactctgt	ctctcaccat	atacaaaagc	3000
	aaatcaaaat	ggatgaaagg	cttaaatcta	aaacctcaaa	cttttgcaact	actaaaagaa	3060
	aacaccggag	aaactctcca	ggacattgga	gtgggcaaa	acttcttgag	taattccctg	3120
50	caggcacagg	caaccaaagc	aaaaacagac	aaatgggcatc	atatcaagtt	aaaaagcttc	3180
	tggccagcaa	aggaacaat	caacaagag	aagagacaac	ccacagaatg	ggagaatata	3240
	tttgcaaaact	attcatctaa	caaggaatta	ataaccagta	tataaagga	gctcaaaacta	3300
	ctctataaga	aaaacaccta	ataagctgat	tttcaaaaat	aagcaaaaga	tctgggtaga	3360
	cattttctcaa	aataagtcac	acaaatggca	aacaggcatc	tgaaaatgtg	ctcaacacca	3420
55	ctgatcatca	gagaaatgca	aatcaaaact	actatgagag	atcatctcat	cccagttaaa	3480
	atgggtttta	ttcaaaagac	aggcaataac	aaatgccagt	gaggatgtgg	ataaaaggaa	3540
	acccttgga	actgttggtg	ggaatggaaa	ttgctaccac	tatggagaa	agtttgaaag	3600
	ttctctcaaaa	aactaaaaat	aaagctacca	tacagcaatc	ccattgctag	gtatatactc	3660
	caaaaaagg	aatcagtgta	tcaacaagct	atctccactc	ccacatttac	tgcagcactg	3720
60	ttcatagcag	ccaaggtttg	gaagcaacct	cagtgtccat	caacagacga	atggaaaaag	3780
	aaaatgtggt	gcacatacac	aatggagtac	tacgcagcca	taaaaaagaa	tgagatcctg	3840
	tcagttgcaa	cagcatgggg	ggcactgggc	agtatgttaa	gtgaaataag	ccaggcacag	3900
	aaagacaaac	ttttcatgtt	ctcccttact	tgtgggagca	aaaattaaaa	caattgacat	3960
	agaaaatagag	gagaatgggtg	gttctagagg	ggtgggggac	aggggtgacta	gagtcaacaa	4020
65	taattttattg	tatgttttaa	aataactaaa	agagtataat	tgggttgttt	gtaacacaaa	4080

5 gaaaggataa atgcttgaag gtgacagata ccccatTTac cctgatgtga ttattacaca 4140
 ttgtatgcct gtatcaaaat atctcatgta tgctatagat ataaacccta ctatatttaa 4200
 aattaaaaat ttaatggcca ggcacgggtg ctcatgtccg taatcccagc actttgggag 4260
 gccgaggcgg gtggatcacc tgagggtcagg agtttgaac cagtctggcc accatgatga 4320
 aacctgtct ctactaaaga taaaaaatt agccaggcgt ggtaggcacat acctgtagtc 4380
 ccaactactc aggaggctga gacaggagaa ttgcttgaac ctggggaggcg gaggttgacg 4440
 tgagccgaga tcatgccact gcactgcagc ctgggtgaca gagcaagact ccatctcaaa 4500
 acaaaaacaa aaaaaagaag attaaaattg taatttttat gtaccgtata aatatatact 4560
 ctactatatt agaagttaa aattaaaaa attataaaag gtaattaac acttaattcta 4620
 10 aaataagaac aatgtatgtg gggtttctag ctctgaaga agtaaaagt atggccacga 4680
 tggcagaaat gtgaggaggg aacagtggaa gttactgttg ttagacgctc atactctctg 4740
 taagtgaact aattttaacc aaagacaggc tgggagaagt taaaggagca ttctataagc 4800
 cctaaaacaa ctgctaataa tgggtaaaag taatctctat taattacca taattacaga 4860
 tatctctaaa atcgagctgc agaattggca cgtctgatca caccgtctc tcatttcagg 4920
 15 tgcctttttt ctgtgtgtct tggagatttt cgattgtgtg ttctgttttg gttaaactta 4980
 atctgtatga atcctgaaac gaaaaatggt ggtgatttcc tccagaagaa ttagagtacc 5040
 tggcaggaag cagggtggctc tgtggacctg agccacttca atcttcaagg gtctctggcc 5100
 aagaccagg tgcaaggcag aggcctgatg acccgaggac aggaagctc ggatgggaag 5160
 gggcgatgag aagcctgcct cgttgtgtgag cagcgcatga agtccctta ttacgcttt 5220
 20 gcaagattg ctctggatac catctggaag aggcggccag cgggaatgca aggagtacga 5280
 agcctcctgc tcaaaaccag gccagcagct atggcgccca cccgggctg tgccagaggg 5340
 agaggagtca aggcacctcg aagtatggct taaatctttt ttccacctga agcagtgaac 5400
 aaggtgtatt ctgagggaag cttaggttag gtgccttctt taaaacagaa agtcatggaa 5460
 gaccccttct caaggaaaa ccagacgccc gctctgcggg catttacctc ttctctctct 5520
 25 ccctctcttg ccctcgcggt ttctgacgg gacagagtga ccccggtgga gcttctccga 5580
 gccgtgctg aggacctctc tgcaaggggc tccacagacc ccccgctgg agagaggagt 5640
 ctgagcctgg cttaataaca aactgggatg tggctggggg cggcacagca cggcgggatt 5700
 caaagactta attccatgag taaattcaac ctttccacat ccgaatggat ttggatttta 5760
 30 tcttaataat ttcttaatt tcataaaata acattcagga ctgcagaaat ccaaggcgt 5820
 aaaacaggaa ctgagctatg ttgtccagg tccaaggact taataaccat gttcagaggg 5880
 atttttcgcc ctgaagtact ttatttgggt ttcataaggt ggcttagggg gcaagggaaa 5940
 gtacacgagg agaggcctgg gcggcagggc tatgagcag gcaggggcac cggggagaga 6000
 gtcccgccc tgggaggctg acagcaggac cactgaccgt cctccctggg agctgccaca 6060
 35 ttgggcaacg cgaaggcggc cactgtgctg gtgactcagg accccatacc ggcttctcgg 6120
 gcccccacac actaacccag gaagtacagg agctctgaac ccgtggaac gaacatgacc 6180
 ctgctctgcc tgcttccctg ggtgggtcaa gggtaatgaa gtggtgtgca ggaaatggcc 6240
 atgtaaatta cagcactctg ctgatgggga ccgttctctc catcattatt catcttcacc 6300
 ccaaggact gaatgattcc agcaacttct tcgggtgtga caagccatga caaaactcag 6360
 40 tacaacacac actcttttac taggccacac gagcacggsc cacaccctg atatatag 6420
 agtccaggag agatgaggct gctttcagcc accaggctgg ggtgacaaca gcggtgaac 6480
 agtctgttcc tctagactag tagaccctgg caggcactcc cccagattct agggcctggt 6540
 tgctgtcttc cgaggcgccc atctgcccgt gagactcagc ctggggtgcc acactgaggc 6600
 cagccctgtc tccacacccc ccgctccag gctcagctt ctccagcagc ttcttaaac 6660
 45 ctgggtgggc cgtgttccag cgtactgtc tcacctgtcc cactgtgtct tgtctcagcg 6720
 acgtagctcg caggttccct cctcacatgg ggtgtctgtc tcttcccca acactcacat 6780
 gcgttgaagg gaggagattc tgcgctccc agactggctc ctctgagcct gaacctggct 6840
 cgtggccccc gatgcagggt cctggcgtcc ggtgcaagc tgacctccat ttccaggcgc 6900
 tccccgtctc ctgtcatctg ccggggcctg ccggtgtgtt ctctgtttc tgtgtctct 6960
 50 tccacgtcca gctgcgtgtg tctctgccc ctagggtctc ggggttttta taggcatagg 7020
 acggggcggt ggtgggcccag ggcgctcttg ggaatgcaa catttgggtg tgaaagtagg 7080
 agtgctgtc ctacactagg tccacgggca caggcctggg gatggagccc ccgcccaggga 7140
 cccgcccctc tctgcccagc actttcctgc cccctcctt ctggaacaca gagtggcagt 7200
 ttccacaagc actaagcatc ctcttcccaa aagaccagc attggcaccc ctggacattt 7260
 55 gcccacagc cctgggaatt cactgacta cgcacatcat gtacacactc ccgtccacga 7320
 ccgacccccc ctgttttatt ttaatagcta caaagcagg aaatccctgc taaaatgtcc 7380
 tttacaaac ttgttaaaaca aacgggtcca tccgcacggt ggacagttcc tcacagtga 7440
 gaggaacatg ccgtttataa agcctgcagg catctcaagg gaattacgct gagtcaaac 7500
 tgccacctcc atgggatagc tacgcaacat gctcaaaaag aaagaatttc acccataggc 7560
 60 aggggagtggt ttaggggggg taaggacggt gggggcgcca gctgggggct actgcacga 7620
 ccttttacta aagccagttt cctggttctg atggtattgg ctcagttatg ggagactaac 7680
 cataggggag tggggatggg ggaacccgga ggctgtgcca tctttgccat gcccgagtg 7740
 cctgggcagg ataatgtct agagatgccc acgtcttgat tccccaaac ctgtggacag 7800
 aacccgccc gcccaggggc ctttgagggt gtgactctcg tgaggacct gaggtctggg 7860
 65 atccttcggg actacctgca gggccgaaaa gtaatccagg ggttctggga agaggcgggc 7920
 aggggggtca gaggggggca gcctcaggac gatggaggca gtcagtctga ggctgaaaag 7980

5 ggagggaggg cctcgagccc aggcctgcaa gcgcctccag aagctggaaa aagcggggaa 8040
 gggaccctcc acggagcctg cagcaggaag gcacggctgg cccctagccc accagggccc 8100
 atcgtggacc tccggcctcc gtgccatagg agggcactcg cgctgccctt ctgacatgaa 8160
 gtgtgtgggg atttgcagaa gcaacaggaa acccatgcac tctgaatcta ggattatttc 8220
 10 aaaaacaaag tttacagaaa catccaagga cagggtctga gtgcctccgg gcaagggcag 8280
 ggcaggcacg agtgatttta tttagctatt ttattttatt tacttacttt ctgagacaga 8340
 gttatgtctt tgttgcccag gctggagtgc agcggcatga tcttggctca ctgcaacctc 8400
 cgctcctctg gttcaagcaa ttctcgtgcc tcagcctccc aagtagctgg gatttcaggc 8460
 gtgcaccacc acacccggct aattttgtat ttttagtaga gatgggcttt caccatgttg 8520
 15 gtcaagctga tctcaaaatc ctgacctcag gtgatccgcc caccctcagc tcccaaatg 8580
 ctgggattac aggcacatgc cactgcacct ggccatttta accattttta aacttcccctg 8640
 ggctcaagtc acacccactg gtaaggagtt catggagttc aatttcccct ttactcagga 8700
 gttaccctcc tttgatattt tctgtaattc ttcgtagact ggggatacac cgtctcttga 8760
 catattcaca gtttctgtga ccacctgtta tccccaggga cccactgcag gggcagctgg 8820
 20 gaggctgcag gcttcaggtc ccagtggggg tgccatctgc cagtagaaac ctgatgtaga 8880
 atcaggcgcg aagtgtggag actgtcctga atctcaatgt ctccagtgtg gctgaaacat 8940
 gtgaaatta aagtcacatc ctctactctt actgggattg agccccctcc ctatcccccc 9000
 ccagggcgag agggagtctc ctactcctg tggagggaag aatgatactt tgttattttt 9060
 cactgctggt actgaatcca ctgtttcatt tgttggtttg tttgttttgt tttgagaggc 9120
 25 ggtttcactc ttgtgtctca ggctggaggg agtgcaatgg cgcgatcttg gcttactgca 9180
 gcctctgctc cccaggttca agtgattctc ctgcttccgc ctcccatttg gctgggatta 9240
 caggcacccg ccaccatgcc cagctaattt tttgtatttt tagtagagac ggggggtggg 9300
 ggggttcacc atgttgccca ggctgtgttc gaacttctga cctcagatga tccacctgcc 9360
 tctgcctcct aaagtgtctg gattacaggt gtgagccacc atgcccagct cagaatttac 9420
 30 tctgtttaga aacatctggg tctgaggtag gaagctcacc ccactcaagt gttgtgtgtg 9480
 tttaaagcaa tgatagaatt tttttattgt tgttagaaca ctcttgatgt ttacactgt 9540
 gatgactaag acatcatcag cttttcaaag acacactaac tgcaccataa atactggggg 9600
 gtcttctggg tatcagcaat ctctcattga tgcggggagg cgtttcctcg ccactgcacat 9660
 35 ggtgttaatt actccagcat aatcttctgc ttccatttct tctcttccct cttttaaaat 9720
 tgtgttttct atgttggtct ctctgcagag aaccagtgtg agctacaact taacttttgt 9780
 tggaaacaaat tttccaaacc gccccttctg cctagtggca gagacaattc acaaacacag 9840
 cccctttaaa aggccttaggg atcactaagg ggatttctag aagagcgacc tgtaattccta 9900
 agtatttaca agacgaggtc aacctccagc gagcgtgaca gcccagggag ggtgcgaggc 9960
 40 ctgttcaaat gctagctcca taaataaagc aatttccctc gccagtttct gaaagtagga 10020
 aaggttacat ttaaggttgc gtttgttagc atttccagtgt ttgcccagct cagctacagc 10080
 atccctgcga ggcctcgga gaccagaag tttctcgccc ccttagatcc aaacttgagc 10140
 aaccggaggt ctggattcct gggaagtcct cagctgtcct gcggtgtgtc cggggcccca 10200
 ggtctggagg ggaccagtgg ccgtgtggct tctactgctg ggctggaaat cgggctcctc 10260
 45 agctctgcag tccgaggctt ggagccaggt gcctggacce cgaggctgcc ctccacctg 10320
 tgcggggcgg atgtgaccag atgttggcct catctgccag acagagtgcg gggggccagg 10380
 gtcaaggccg ttgtggctgg tgtgaggcgc ccggtgcgcg gccagcagga gcgcctcgct 10440
 ccatttccca ccccttctcg acgggaccgc cccgggtgggt gattaaacaga tttgggggtg 10500
 tttgctcatg gtggggaccc ctgcgccctc gagaacctgc aaagagaaat gacgggctcg 10560
 50 tgtcaaggag cccaagtgcg ggggaagtgt tgcaggaggg cactccggga ggtcccgcgt 10620
 gcccgctccag ggagcaatgc gtccctcggtt tgcctccag ccgcgtctac gcgcctccgt 10680
 cctccccctc acgtccggca ttctgtgtgc ccggagcccc acgccccgcg tccggacctg 10740
 gaggcagccc tgggtctccg gatcaggcca gcggccaaag ggtcgccgca cgcacctgtt 10800
 cccagggcct ccacatcatg gcccctccct cgggttacc cccagcctag gccgattcga 10860
 cctctctccg ctggggccct cgctggcgct cctgcacctt gggagcgca gggggcgcg 10920
 55 ggcggggaaag cgcggcccaag acccccgggt ccgcccggag cagctgcgct gtcggggcca 10980
 ggccgggctc ccagtggatt cgcgggcaca gacgcccagg accgcgtccc ccacgtggcg 11040
 gagggactgg ggaccggggc acccgtcctg ccccttcacc ttccagctcc gcctcctccg 11100
 cgcggaaccc gcccggtccc gacccctccc ggggtccccg cccagccccc tccgggccc 11160
 cccagccctt ccccttctct tccgcggccc cgcctctccc tccggcgcg agtttcagge 11220
 60 agcgtgtcgt cctgtgtgc acgtgggaag ccctggcccc ggccaccccc gcgatg 11276

 <210> 4
 <211> 104
 <212> DNA
 60 <213> Homo sapiens

 <400> 4
 gtgggcctcc cgggggtcgg cgtccggctg ggggttgagg cgccggggg gaaccagcga 60
 catgcggaga gcagcgcagg cgactcagg cgcttcccc gcag 104
 65

7 / 18

<210> 5
<211> 8616
<212> DNA
<213> Homo sapiens

5

<400> 5

gtgaggaggt ggtggccgtc gagggcccag gccccagagc tgaatgcagt aggggctcag 60
aaaagggggc aggcagagcc ctggctcctcc tgtctccatc gtcacgtggg cacacgtggc 120
ttttcgtcta ggacgtcgag tggacacggg gatctctgcc tctgctctcc ctccgtgcca 180
10 gtttcgataa acctacgagg ttacacctca cgttttgatg gacacgcggg ttccaggcgc 240
cgaggccaga gcagtgaaca gagggaggtg ggcgcggcag tggagccggg ttgcccggca 300
tggggagaag tgtctggaag cacagacgct ctggcgaggg tgcctgcagg ttacctataa 360
tctcttctgc aatttcaagg gtgggaatga gaggtgggga cgagaacccc ctcttctctg 420
gggtgggagg taagggtttt gcaggtgcac gtggtcagcc aatatgcagg tttgtgttta 480
15 agatttaatt gtgtgttgac gggcagggtg ggtggctcac gccggtaatc ccagcacttt 540
gggaagctga ggcaggtgga tcacctgagg tcaggagttt gagaccagcc tgaccaacat 600
ggtgaaaccc tatctgtact aaaaatacaa aaattagctg ggcattggtg tgtgtgcctg 660
taatcccagc tacttgggag gctgaggcag gagaatcact tgaaccagag aggcggaggc 720
tgacgtgagc tgagattgtg ccattgtact ccagcctggg cgacaagagt gaaactctgt 780
20 ctttaaaaaa aaaaagtgtt cgttgattgt gccaggacag ggtagaggga gggagataag 840
actgttctcc agcacagatc ctgggtcccat ctttaggtat gaagagggcc acatgggagc 900
agaggacagc agatggctcc acctgctgag gaaggagacg tgtttgtggg tgttcagggg 960
atgggtctgc tgggcccctg cgtgtcccca cctgttttt ctggatttga tgttgaggaa 1020
cctccgctcc agcccccttt tggctcccag tgcctccagg cctaccctg gcagctagaa 1080
25 gaagtcccca tttcaccccc tccccacaaa ctcccaagac atgtaagact tccggccatg 1140
cagacaagga gggtagacct cttggggctc ttttttttct ttttttcttt ttatgggtgc 1200
aaaagtcata taacatgaga ttggcactcc taacaccgtt ttctgtgtac agtgcagaa 1260
tgctaactcg gcggtgttta cagcagggtg cttgaaatgc tgcgtcttgc gtgactggaa 1320
gtccctaccc atcgaacggc agctgcctca cactgtctgc ggctcagggt gaccacggc 1380
30 agtcagataa gcgtcatgca acccagtttt gctttttgtg ctccagcttc cttcgttag 1440
gagagttaga gttctctgat caggactctg cctgtcattg ctgttctctg acttcagatg 1500
aggtcacaat ctgccccctg cttatgcagg gagttagggc tgggtccccg gtgtccctgt 1560
cacgtgcagg gtgagttagg cgttgccccc aggtgtccct gtcacgtgta ggtgagtg 1620
ggcgcggccc cgggtgttcc ctgtcccgtg cagcgtgatt gaggtgtggc ccccggtgt 1680
35 cctgtcacg tgtagggtga gtgaggcgcc atccccgggt gtccctgtca cgtgtagggt 1740
gagttagggc tggccccggg gtgtccctgt cccgtgcagg gtgagttagg cactgtcccc 1800
gggtgtccct gtcacgtgca ggtgagtga ggcgcgggtc cgggtgttcc ctctcagggt 1860
tagggtgagt gaggcgcggc cccagggtgt cctgtcacg tgtagggtga gtgaggcacc 1920
gtccctgggt gtccctccca ggtatagggt gagttaggca ctgtccccg gtgtccctgt 1980
40 cacgtgcagg gtgagttagg cgcggccccc ggggtgtccct ctccaggtgca ggtgagtg 2040
ggcgtgttcc ctgggtgtcc ctgtctctgt tagggtgagt gaggtctgt cccaggtgt 2100
ccttggcgtt tgcctcactg agcttctctc tgaatgtttg ctcttcttat agccacagct 2160
gcgcgggtt cccattgcct ggttagatgg tgcaggcgca gtgctgggtc ccaagcctat 2220
cttttctgat gctcggtctt tcttgggtac ctctccgttc cattttgtta cggggacacg 2280
45 ggactgcagg ctctcgcttc ccgcgtgcca ggcactgcag ccacagcttc aggtccgctt 2340
gcctctgttg ggcctggctt gctcaccacg tgcggccac atgcatgctg ccaatactcc 2400
tctcccagct tgtctcatgc cagggtgga cctggggctg cctgtgtctg ctgccacgtg 2460
ttgctggaga catcccagaa aggggtctct gtgccctgaa ggaaagcaag tccccacagc 2520
ccctcactt gtctgtttt ctcccaagct gccctctctg ttggccccc ttgggtgggtg 2580
50 gcaacgctt tcaccttatt ctgggcacct gccgtcatt gcttaggtg ggctctgcct 2640
ccagtcgccc cctcacatgg attgacgtcc agccacaggt tggagtgtct ctgtctgtct 2700
cctgtctga gacccacgtg gagggccggg gtctccgcca gccctctgta gacttccctc 2760
ttgggtctta gttttgaatt tcactgattt acctctgacg tttctatctc tccattgtat 2820
55 gcttttctt ggttttattt ttcttctctt tcttagcttc ttagttagt catgctcttc 2880
cctctaagtg ctgccttacc tgcacctgt gttttgatgt gaagtaatct caacatcagc 2940
cactttcaag tgttcttaaa atacttcaaa gtgttaatac ttcttttaag tattcttatt 3000
ctgtgatttt tttctttgtg cagcgtgtgt ttgacgtga aatcattttg atatcagtg 3060
cttttaagta tcttttagct tattctgtga tttctttgag cagttagtta ttgaaacct 3120
gtttatgttc aagatatgta gagtatcaag atacgtagag tattttaagt tatcatttta 3180
60 ttattgattt ctaactcagt tgtgtagtgg tctgtataat accaattatt tgaagttctg 3240
ggagccttgc tttgtgatct agtgtgtgca tgggtttccag aactgtccat tgtaaatttg 3300
acatcctgtc aatagtgggc atgcatgttc actatatcca gcttattaag gtccagtgca 3360
aagcttctgt ctccctctag atgcatgaaa ttccaagaag gaggcataag tccctcact 3420
65 gggggatggg tctgttctat tcttctctgt tggtagcatt tatgtgaggc attgttaggt 3480
gcatgcacgc ggtagaattt ttatcttctt gatgagtga tcttttgag acttctatgt 3540

ctctagtaaat cttagtaattc ttttttttaa ttgctcttag tactgccaca ctgggcttct 3600
 tttagattagt attttcctgc tgtgtctgtt ttctgccttt aatttatata tatatatata 3660
 tttttttttt tttttgagaca gagtcttggg ctgtcgccca gggtagtgag agtgggtgtga 3720
 5 tccacaggtea gtgtaacttt taccttctgg cctgagccgt cctctcacct cagcctcctg 3780
 agtagctgga actgcagaca cgcaccgcta cactggcta atttttaaat tttttctgga 3840
 gacaggggtct tgctgtgttg cccaggctgg tctcaaaact ttggactcaa gggatccatc 3900
 tacctcgggt tcccaaaagt ctgaattaca ggcagagcc accatgtctg gcctaatttt 3960
 caacactttt atattcttat agtgtgggta tgtcctgtta acagcatgta ggtgaatttc 4020
 10 caatccagtc tgacagtcgt tgtttaactg gataacctga tttattttca tttttttgtc 4080
 actagagacc cgcctgggtg actctgattc tccacttgcc tgttgcatgt cctcgttccc 4140
 ttgtttctca ccactcttg ggttgccatg tgcgtttcct gccgagtggt tgttgatcct 4200
 ctggttgctt cctggtcact gggcatttgc tttattttct tttgtcttag tgttaccctc 4260
 tgatcttttt attgtcgttg tttgtctttg tttatttgaga cagtctcact ctgtcaccca 4320
 ggctggagtg taatggcaca atctcgggct actgcaacct ctgcctcctc ggttcaagca 4380
 15 gtcttcattc ctcaacctca tgagtagctg ggattacagg cgccaccac cagcctcctg 4440
 taatttttgt atttttagta gagataggct ttcaccatgt tggccagggt ggtctcaaac 4500
 tctgacctc aagtgatctg cccgccttgg cctcccacag tgcctgggatt acaggtgcaa 4560
 gccaccgtgc cgggcatacc ttgatctttt aaaaatgaagt ctgaaacatt gctacccttg 4620
 20 tccctgagcaa taagaccctt agtgtatttt agctctggcc acccccacgc ctgtgtgctg 4680
 tttttccctg tgacttagtt ctatctcagg catcttgaca cccccacaag ctaagcatta 4740
 ttaattattgt tttccgtgtt gagtgtttct gtacttttgc ccccgccctg cttttctctc 4800
 ttgtttcccc gtctgtcttc tgtctcaggc cgcgcgtctg gggctcccct ccttgcctt 4860
 tgcgtgggtc ttctgtcttg ttattgtctg taaaccccag ctttacctgt gctggcctcc 4920
 atggcatcta gcgagctccg ggggacctctg cttatgatgc acagatgaag atgtggagac 4980
 25 tccagaggag ggcggtcatc ttggcccgtg agtgtctgga gcaccacgtg gccagcgttc 5040
 cttagccagt gaggtagacg aacgtccgct cggcctgggt tcagcctgga aaaccccagg 5100
 catgtcgggg tctgtgtggt cgcggtgtgc gaggttgaaa tcgcgcaaac ctgcggtgtg 5160
 gcgcagctc tgacggtgct gcctggcggt gaggatgtctg cttcctcctc tctgcttggg 5220
 aaccaggaaca aaggatgagg ctccgagccg ttgtcgccca acaggagcat gacgtgagcc 5280
 30 atgtggataa ttttaaaatt tctaggctgg gcgcggtggc tcacgcctgt aatcccagca 5340
 ctttggggag ccaaggcggtg tggatcacga ggtcaggagg tcgagaccat cctggccaac 5400
 atgatgaaac cccatctgta ctaaaaacac aaaaattagc tgggcgtggt ggcggtgctc 5460
 tgaatccca gctactcggg aggetgagggc aggagaattg cttgaacctg ggagtggaa 5520
 35 gttgcagtga gccgacattg caccactgca ctccagcctg gcaacacagc gagactctgt 5580
 ctcaaaaaaa aaaaaaaaaa aaaaaaaaaa aattctagta gccacattaa aaaaatgaaa 5640
 aagaaaaagg gaaattaatg taataataga ttttactgaa gccagcatg tccacacctc 5700
 atcatttttag ggtgttattg gtgggagcat cactcacagg acatttgaca ttttttagc 5760
 tttgtctgcy ggtaccctg ttaggtccc gtgcgtggcc atctcggcct ggacctgctg 5820
 40 ggccttccat ggccatggct gttgtaccag atggtgcagg tccgggatga ggtcgccagg 5880
 ccttcagtga gctggatgt cagtgtccgg atggtgcag tctgggatga ggtcgccagg 5940
 cctctgtgtg agctggatgt gtggtgtctg gatggtgagc gtacgggggt aggtctccag 6000
 gccctcgggt agctggaggt atggagtcg gatgatgag gtccgggggt aggtcgccag 6060
 45 ggcctcgtgt gagctggatg tgtggtgtct ggatggtgca ggtcagggggt gaggctccca 6120
 ggccctcgggt aagctggagg tatggagtcc ggatgatgca ggtccgggggt gaggctccca 6180
 ggcctcgtgt tgagctggat gtgtggtgtc tggatgggtg aggtctgggg tgaggtcacc 6240
 aggcctcgtg gtgagctggg tgtgcggtgt ctggatgggt caggctctga gtgaggtcgc 6300
 cagacggtgc cagaccatgc ggtgagctgg atatgcggtg tccggatggt gcaggtctg 6360
 50 ggtgaggttg ccaggccctg ctgtgagttg gatgtgggtg gtccggatgc tgcaggtcgc 6420
 ggttgagggt accaggccct gctgtgagct ggatgtgtgg tgtctggatg gtgaggtctc 6480
 ggggtgaggg tgcgccaggc cctgcttgtg agctggatgt tgggtgtctg gatggtgcag 6540
 gtcttgagtg aggtcgccag gccctcgggt agctggatgt gcagtgtcca gatggtgcag 6600
 gtccgggggt aggtcgccag accctcgggt gagctggatg tgcggtgtct ggatggtgca 6660
 55 ggtctggagt gaggctcgcc agccctcgggt gagctggatg tatggagtcc ggatggtgccc 6720
 ggtccgggggt gaggctcgcc gaccctcgtg tgagctggat gtgcggtgtc tggatggtac 6780
 aggtctggag tgaggtcgcc agaccctgct gtgagctgga tatgcggtgt ccggatggtg 6840
 caggctcagg gtgaggtctc caggccctcg gtgagctgga ggtatggagt ccggatgatg 6900
 caggctcagg gtgaggtcgc caggccctgc tgtgaactgg atgtgcggcg tctggatggt 6960
 gcaggtctgg ggtgtggtcg ccaggccctc ggtgagctgg aggtatggag tccggatgat 7020
 60 gcaggtctcg ggtgaggtcg ccaggccctc ctgtgagctg gatgtgcggc gtctggatcg 7080
 tgcaggtctg ggtgtggtgc gccaggccct cgtgagctg gaggatgga gtccggatga 7140
 tgcaggtctg ggtgtgaggt gccaggccct gctgtgagct ggatgtgctg tatccggatg 7200
 gtgcagctcg ggtgtgaggt gccaggccct gctgtgagct ggatgtgctg tatccggatg 7260
 gtgcaggtct ggtgtgaggt caccaggccc tgcggtgagc tgggtgtgct gtgtccgggt 7320
 65 gctgcaggtc cgggtgaggt tgcagggccc ctgcggtgagc tggatgtgct gtgtcccggt 7380
 gtcggatggt tgcaggtcca ggtgtgaggt gctaggccct tgggtgggtg gatgtgcctg 7440

gtccggatgg tgcaggctcg ggggtgaggtc gccaggccctt tggtagctg gatgtgcggg 7500
gtctgcatgg tgcaggctcg ggggtgaggtc gccaggccctt tggtagctg gatgtgcggg 7560
gtccggatgg tgcaggctcg gctgtgaggtc gccaggccctt gctgtgagct ggatgtgcgg 7620
tctctggatg gtgcaggctc ggggtgaggt agccaaggcc ttcggtagc tggatgtggg 7680
5 gtgtccggat ggtgcaggct cggggtgagg tgcaggccctt ctgaggtag ctggatatgc 7740
gggtgtccga tgggtgcagg cggggtgagg gtcaccaggc cctgaggta gctggatgtg 7800
cgggtctctg atgggtgcagg tccgggggtga ggtcgccagg cctgtctgtg agctggatgt 7860
gctgtatccg gatgggtgca gtcgggggtg aggtcgccag gccctgcagt gagctggatg 7920
tctgttatcc ggatgggtgca ggtctggcgt gaggtcgcca ggcctcgagg ttagctggat 7980
10 atgagggtgc ggatgggtgca ggtccggggg gaggtcacca ggcctcgagg ttagctggat 8040
gtgagggtgc cggatgggtc aggtctgggg tgagggtcgc aggcctcgct gtgagctgga 8100
tctgctgtat ccggatgggt cagggtccggg gtgagggtgc caggccctgc ggtgagctgg 8160
atgtgctgta tccggatggg gcaggctctg cgtgagggtc ccaggccctg cggtagctg 8220
gatgtgcagt gtacggatgg tgcaggctcg ggggtgaggtc gccaggccctt ggggtgggct 8280
15 tggatgtgtg tgtctggatg gtgcaggctc ggggtgaggt cgcaggccct tgcggtagc 8340
tggatgtgtg gtgtctggat gctgcaggct cggggtgaggt tgcaggccct ctggtgagc 8400
tggatgtgtg gtgtccccgt gtccgaatgg tgcagggtcca gggtagaggtc gccaggccct 8460
tggtagctg gatgtgcccgt gtccggatgg tgcaggctcg gggtagaggtc gccaggccct 8520
20 tggtagctg gatgtgcggg gtccggatgg tgcaggctcg gggtagaggtc gccaggccct 8580
cggtagctg gatgtggcat gtccttctcg ttttaag 8616

<210> 6
<211> 2089
<212> DNA
25 <213> Homo sapiens

<400> 6
gtactgtatc cccacgccag gcctctgctt ctggaagtc tggaaacca gcccgccctc 60
agcatgcgcc tgtctccact tgcctgtgct tccctggctg tgcagctctg ggtggggagc 120
30 caggggcccc gtccacaggc tgggtccagt ggattctgtg caaggctctg actgcttga 180
gctcacgttc tcttacttgt aaaatcagg gttgtgcca agtggctctt agggtttga 240
aagcagaagg gatataaatt agatggaaac actaccacta gcctccttgc ctttcccttg 300
gatgtgggtc tgattctctc tctctttttt ttttctttt tgagatggag tctcactctg 360
ttgccaggc tggagtgcag tggcataatc ttggctcact gcaacctcca cctcctgggt 420
35 ttaagcgatt caccagcctc agcctcctaa gtactctgga ttacaggcac ctgccaccac 480
gcctggctaa tttttgtact tttaggagag acgggggttc accatgttgg ccaggctggg 540
ctcgaaactc tgacctcagg tgatccaccc accttggcct cccaaagtgc tgggtttaca 600
ggctaagcca cgtgcccag ccccgattc tcttttaatt catgctgttc tgatgaatc 660
ttcaatctat tggatttagg tcatgagagg ataaaatccc acccacttgg cgactcactg 720
40 cagggagcac ctgtgcaggg agcacctggg gataggagag ttccaccatg agctaacttc 780
taggtggctg catttgaatg gctgtgagat tttgtctgca atgttcggct gatgagagt 840
tgagattgtg acagattcaa gctggatttg catcagtgag ggacgggagc gctgggtctg 900
gagatgccag cctggctgag cccaggccat ggtattagct tctccgtgtc ccgcccaggc 960
tgactgtgga gggcttttagt cagaagatca gggcttcccc agctccctg cactctgag 1020
45 tccctggggg gccttgtgac acccctatgc ccaaatcagg atgtctgag agggagctgg 1080
cagcagacct cgtcagaggg aacacagcct ctgggctggg gaccccgagc tgggtgctggg 1140
gccatttcc tgcactctgg gagggtcag ggccttccct gtgggaacaa gttatacac 1200
aatgcacctt acttagactt tacacgtatt taatgggtgt cgacccaaca tggctatttg 1260
accagtattt tggaaagaat ttaattgggg tgaccggaag gagcagacag acgtgggtgg 1320
50 ccccaagatg ctcttctgca ctactgggac tgtgttctg cctggggggc cttggaggcc 1380
cctcctccct ggacagggtg ccgtgccttt tctactctgc tgggctctgc gctgctgggt 1440
agggcaccag ctccggagca cccgaggccc cagtgtccac ggagtgccag gctgtcagcc 1500
acagatgccc aggtccaggg gtggcgcgtc cagccccctg gcccccatgg ggtgttttg 1560
55 gggaaaaggc caagggcaga ggtgtcagga gactgggtgg ctcatgagag ctgattctgc 1620
tcttggctg agctgcccgt agcagcctct cccgcccctt ccacttgaag ggtgtgggt 1680
ctttctacct gggggtcctg cctggggcca gccttgggt accccagtgg ctgtaccaga 1740
gggacaggca tctgtgtgag aggggcatgg gttcacgtgg cccagatgc agcctgggag 1800
caggctcccc ggtgctgag gtgggacagt caccctgggg gttgaccgcc ggactgggag 1860
tcccagggtg tgactatagg accaggtgtc cagggtcccc gcaagtagag gggctctcag 1920
60 aggcgtctgg ctggcatggg tggacgtggc cccgggcatg gccttcagcg tgtgctgccc 1980
tgggtgcccgt gagccctcac tgagtcgggt ggggcttctg gcttcccctg agcttcccc 2040
tagtctgttg tctggctgag caagcctcct gaggggctct ctattgagc 2089

10

15

25

30

40

45

6

6

agctttattg agggagaccat atcttccttt gaactatggt cgggtttata gtaagtcagg 180
 ggtgtggagg cctcccttgg gctccctgtt ctgtttcttc cactctgggg tcgtgtgggt 240
 cctgtgtgtg tgtgtggcgg gtgggcagggt ctccaggcc tccttgtgtt cattggcctg 300
 gatgtggccc tgggtacgct ccgtccctgg aattcccctg cgagttggag gctttctttc 360
 5 tttctttttt tctttctttt tttttttttt tgataacaga gtctcgctct tttttgcca 420
 ggctggagtg gtttggcgtg atcttggctc actgcaacct gtgcttcctg agttcaagca 480
 attctcttgc ctccagcctc caagtagctg gaattatagg cgccaccac catgtgact 540
 aatttttcta atttttagtag agacgaggtt tctccatgtt ggccaggctg gtctcgaact 600
 10 cctgacctca ggtgatcctc ccacctcggt cctccaaagt gctgggatga caggtgtgaa 660
 ccgcccggcc cggccgagac tcgcttcctg cagcttccgt gagatctgca gcgatagctg 720
 cctgcagcct tgggtgtgac aacctccgtt tctcttctcc aggtctcgct aggggtcttt 780
 ccatttcatg actctcttca cagaagagtt tcacgtgtgc tgatttcccg gctgtttctt 840
 gcgtaattgg tgtctgctgt ttatcgatgg cctccttcca tttctttag gctttgttta 900
 ttgtgtttt tccggctcct tgaaggaaaa gtctcgatta tggatgtttg aactttcttt 960
 15 tctaacaag catctgaagt tgccgttttc cctctaaagc agggatcccc aggccctgg 1020
 ctgtggagtg gcaccgggtt ggggctgtt aggaaccggg cgccagcggg gaggctagg 1080
 ggggtgtggg gagccagcgt tcccgctga gcccgcccc tctcagatca gcagtggcat 1140
 gcgggtgtca gaggcgaca caccctactg agaactgtgc gtgagagggg tctagattct 1200
 20 gtgtcctta tgggaatcta atgctgatg atctgaggtg gaaccgtttg ctcccaaaa 1260
 catccccctt cccactgctg tctgtggaa aaatcgtctt ccacgaaacc agtccctgg 1320
 accacaatgg ttggggaccc tgtgtctaaag acctgtctca gcagcctctc gtcagtgtt 1380
 atatatggc tttctgtgtg tgagtcacaga ataattacgg atttctgtga tgctttccg 1440
 cgacctcaga cccatgggct atttggggc gtgttgctg ctctgggtt gggaaagggtg 1500
 25 caggcccat gtaccttctt gtactgctt tccagggtg tttcagggtg tgaatcgtac 1560
 tcgagtgggt tttagccac ggccttgcg ccagctcctg ggggctggg aacatgtga 1620
 agcacagagt caccgtgcg gtcttttgat gctcacaag ctcgaggcct cctgtgtccg 1680
 tgttagtgtg tgtcacgtgc ctgtcacat cctgtcttg ggacgcaggg gcttagcagg 1740
 tcccgtagta aatgacaagc gtctggggg agtctgcaga ataggaggtg ggggtgccc 1800
 30 tctctctccc gcttctcag actcttctc tgctgtgct gtggtgcac ctgcatccct 1860
 gcaatccctc cagcactggg ctggagaggc cgggagctc gagtgcact tgtgccact 1920
 gactgtggat ggcagtcggt caggggggtc tgatgtgtg tgactgtgga tggcggttgg 1980
 tcacaggggt ctgagtgtg gtgactgtg atggcggtcg tggggtctga tgtggtgact 2040
 35 tggatggcg gtctgtgggt ctgagtgtg gtgactgtg atggcggtcg tggggtctga tgtggtgact 2100
 tgtgtgact gtggatggcg gtctgtgggt ctgagtgtg gactgtggat ggcggtcgtg 2160
 ggggtctgat tgggtactgt ggtggcaggt cgtggggtct gatgtgtgt gactgtggat 2220
 ggcggtcgtg ggggtcgtg tgggtactgt ggtggcaggt cgtggggtct gatgtgtgt 2280
 gactgtggat ggcggtcgtg ggggtctgat tgtgtgact gtggatggcg gtctgggggt 2340
 40 ctgagtgtg gtgactgtg atggcggtcg tggggtctga tgtgtgtga ctgtggatgg 2400
 cgggtcgtgg gtctgatgtg gtgactgtg atggcggtcg tggggtctga tgtgtgtga 2460
 ctgtggatgg tgatcgggtc caggggtctg atgtgtgggt actgtggatg gcggtcgtg 2520
 ggtctgatgt gtggtgactg tggatgggtga tccgtcacag ggggtctgat tgtgtgact 2580
 gtggatggcg gtctgtgggt ctgagtgtg gtgactgtg atggcggttg gtcccgggg 2640
 45 tctgatgtgt ggtgactgtg gatggcgatc ggtcacagg gtctgatgtg tgggtactgt 2700
 ggtggcggt cgtggggtct gatgtgtgt gactgtggat ggcggtcgtg ggggtctgat 2760
 tgtgtgact gtggatggcg gtctgtgggt ctgagtgtg gactgtggat ggcggtcgtg 2820
 ggggtctgat tgggtactgt ggtggcggt cgtggggtct gatgtgtgt gactgtggat 2880
 50 ggcggttgggt cccgggggtc tgatgtgtg tgactgtgga tggcggtcgt ggggtctgat 2940
 gtggtgactg tggatggcag tctgtgggtc tgatgtgtg tgactgtgga tggcggtcgt 3000
 ggggtctgat gtgtgtgac tgtggatggc ggtcgtgggt tctgatgtg ggtgactgtg 3060
 gatggcggtc gtggggtctg atgtgtgtg actgtggatg gcggtcgtg ggtctgatgt 3120
 ggtgactgtg gatggcggtc gtggggtctg atgtgtgtg actgtggatg gtgactcgtc 3180
 acaggggtct gatgtgtgt gactgtggat ggcggtcgtg ggggtctgat tgtgtgact 3240
 55 tgggatggcg gtctgtgggt ctgatgtgt gactgtggat ggcggtcgtg ggggtctgat 3300
 tgtgtgact gtggatggcg gtctgtgggt ctgatgtgt gtgactgtg atggcagtcg 3360
 gtccagggg tctgatgtgt ggtgactgtg gatggcggtc gtggggtctg atgtgtgtg 3420
 actgtggatg ggcgtcgtg ggtctgatgt gtggtgactg tggatggcg tctgtgggtc 3480
 60 tgatgtgtg tgactgtgga tggcggtcgt ggggtctgat tgggtgactg tggatgggtga 3540
 tccgtcacag ggggtctgat tgtgttagct gcaggtggag tcccaggtgt gtctgtagct 3600
 accttgcgtc ctccggcccc cggccccctt tccccaaaca gaagcttccc aggcgtctc 3660
 tgggttcat cccggcatcg ggttggccg cagggtccaca cgtcctgac ggaagaaaca 3720
 agtgcacagc tctggccggg gcaggccaca tttgtggctc atgcccctc ctctgcccgc 3780
 ag 3782

12 / 18

<210> 11
<211> 980
<212> DNA
<213> Homo sapiens

5

<400> 11
gtctcgggcac tgccttcgag ggttgggcac ggactcccag cagtgggtcc tccccgggc 60
aatcactggg ctcacgaccg gacagactgt tggccctggg gggcagtggg ggaatgagc 120
tgtgatgggg gcatgatgag ctgtgtgcct tggcgaaatc tgagctgggc catgccaggc 180
10 tgcgacagct gctgcattca ggcacctgct cacgtttgac tgcgcgccct ctctccagtt 240
ccgcagtgcc ttgtttcatg atttgcctaa tgccttctct gccagttttg atctttgaggc 300
caaaggaag gtgtccccc cctttaggag ggcaggccat gtctgagccg tgccttgcgc 360
agctggcccc tcagtgtcgg gtctgaggcc aaaggaacg tgtcccccct cttaggaggga 420
cgggcccgtgt ttgagccacg ccccgctgag cgggcccctc agtgcctgggt ctgtccacgt 480
15 gggccctgtg cccctttgcag atgtggtctg tccacgtggc cctgtggctc ttgacagatg 540
cctgttagca cttgctcggc tctaggggac agtcgtgtcc accgcatgag gctcagagac 600
ctctgggcga atttcccttg cttccagggt gggggtggag gtggcctggg ctgctgggac 660
ccagaccctg tgcggcgag ctgggcagca actcctggat cacatatgcc atccggggcca 720
cgggtgggctg tgtgggtgtg agccacagct gacccacagg tggccagag gagacgttct 780
20 gtgtcacaca ctctgcctaa gcccattgtg gtctgcagag actcggcccg gccagccac 840
gatggccctg cattccagcc cagccccgca cttcatcaca aacactgacc ccaaaaggga 900
cggaggggtct tggccacgtg gtcttcgctg tctcagcacc caccgggtca ctcccatgtg 960
tctcccgtct gcttttcgag 980

25

<210> 12
<211> 2485
<212> DNA
<213> Homo sapiens

30

<400> 12
gtgagtcagg tggccagggt ccattgccct gcgggtggct gggcgggctg gcagggtctc 60
tgctcacctc tctcctgccc cttccccact gnccttctgc cgggggccac cagagtctcc 120
tttcttgccc cccgccccct ccggtcctcg ggtgcaggc tcccagggcc ccggaacat 180
35 ggctcggctt gcggcagccg gagcgagca ggtgccacac gaggcctgga aatggcaagc 240
gggtgtgtga gttgctcctg cgtggaggac gaggggcggg ggggtgtgtc gggctcagggt 300
tgcgcccagc gtttgagcct cagccttctc agctccaaat tactactgac gctggacacc 360
cggctctcac acgcttctat ctctctctcc cgatacaaaa ggattttatc cgattctcat 420
tccctgctcc gtctgtgac ccccgcgagg gcgcgggctc tctctctctg tactagattt 480
40 cccatctgga aagtgcgggg ttgaccgtgt agtttgcctc tctcgggggg cctgtggtg 540
ccatggggca ggcggcctgg gagagctgcc gtcacacagc cactgggtga gccacactca 600
cgggtgtaga gccacagtgc ctggtgccac atcacgtcct ctggatttta agtaaaacca 660
cacacctccc ggcaggcatc tgcctgcgac cctgtgtgtg cctggggaga gtggtagcac 720
ggaggaaatt cgtgcacact caaggtcatc agcaaggtca tccgcagtc ggtggaacgt 780
ggaggcctct ctctgggacg gtctccagcg gataaaggac tgtgcacagc ttcggaagct 840
45 tttattttaa aatataacta ttaattattg cattataagt aatcactaat ggtatcagca 900
attataatat ttattaaagt ataattagaa atattaagta gtacacacgt tctggaaaaa 960
cacaaattgc acatggcagc agagtgaatt ttggccgagg gacacgtgtg cacatgtgtg 1020
taagcgcccc ccaggcccc agaattcgtc gacaaagtca cctccccaga gaagccacca 1080
cgggcccctc tctgtgtcgt gaattttatt aagatggatc aagtcacgta ccgtccacgt 1140
50 gtggcagggc tttggggaat gtgaggtgat gactgcgtcc tcatgccctg acagacagga 1200
ggtgactgtg tctgtcctgt ccttaggaca cggacaggcc cgaagctcta gtccccatcg 1260
tgtgtccagt tggcctctga ataaaaacgt cttcaaaacc tgttgcccca aaaactaaga 1320
acagagagag tttcccatcc catgtgctca caggggcgta tctgcttgcg ttgactcgtc 1380
gggctggccg gactcctaga gttggtgcgt gtgcttctgt gcaaaaagtg cagtctctct 1440
55 gcccactact gtgatattcg caccagcaag gaaagcctct tttcttttct tttttttt 1500
ttttttgaga cggaacgtca ctgttgcctg cctgggcttg agtgcagtgg cgcgactctca 1560
actcactgca acctcgcct cccgggttcc agcatttctc ctgcccagc ctcccagca 1620
gctgagatta caggcaccca cccctgcgc ctggctaatt tttgtatttt tagtagagag 1680
gggtttttgc catgttggcc aggtgtgtct cgaactcctg acctcagggt atcccccac 1740
60 ctcggcctcc caaagtgtg ggattacagg tgtgagccat cagccccagc cggaaagcct 1800
ctttttaagg tgaccaccta tagcgttctc cgaaaaatac aggtcctgtt ttgacagtag 1860
ctgtcaagcg tctcttagca acaggagtgg cgtcctgtgg gctctgggga tggctgaggg 1920
tcgctgggca gccatgcctt ctgtgtgcac ctttaggttc cagggggcta tctgtctctc 1980
actgtttgtc tgaaaaacga cccctggcat ccttgtttgg agagtctctg cttctcgttg 2040
65 gtcatgctga aactaggggc aaggttgtat ccgttggcgc gcagcggtca catgtagggt 2100

13 / 18

```

catgagtcctt tcaccgtgga caaatccctt gaaaaaaaaa aaaggagctcc ggtaaagcat 2160
tcattccggg tcaagtgtct ggttctgtga ataaactcta agatttaaga aaccttaatg 2220
aaagaaaacc ttgatgatc agagcaagga tgtggtcaca cctgtggctg gatctgtttc 2280
agccgccccca gtgcatgggt agagtgggga gcagggtattg tttgttcaga ggtctcatct 2340
5 ggtagtgttc tgagggtgtt gccggctgaa tggtagacgt gtcgtttgtg tgtatgaggt 2400
tctgtgtctg tgtgtggctc ggtttgagt tacgcatgtc cagcacatgc cctgcccgtc 2460
tctcacctgt gtcttcccgc cccag 2485

<210> 13
10 <211> 1984
    <212> DNA
    <213> Homo sapiens

    <400> 13
15 gtgaggcctc ctcttcccca ggggggcttg ggtgggggtt gatttgcttt tgaatgcattc 60
    agtggttaata ttcttggtgc tctggagacc atgactgtct tgtcttgagg aaccagacaa 120
    ggttgcagcc ccttcttggt atgaagccgc acgggagggg ttgcacagcc tgaggactgc 180
    gggctccacg caggctctgt ccagcggcca tgtccagagg cctcagggct cagcaggcgg 240
    gagggccgct gccctgcatt atgagcatgt gaattcaaca ccgaggaagc acaccagctt 300
20 ctgtcacgtc acccaggttc cgttagggtc cttggggaga tggggctggt gcagcctgag 360
    gccccacatc tcccagcagg ccctcgacag gtggcctgga ctgggcgcct cttcagccca 420
    ttgccatcc cacttgcatg gggctctacac ccaaggacgc acacacctaa atatcgtgcc 480
    aacctaatgt ggttcaactc agctggcttt tattgacagc agttactttt ttttttttaa 540
    tacttttaagt tctagggtag atgtgcacga cgtgcagggt agttacatat gtatacatgt 600
25 gccatggttg tgtgtgcac ccattaactc atcatttaca ttaggtatat ctccaatatg 660
    tatccctccc cactcccccc atcccatgac aggccttggt gtgtgatgtt cccacccttg 720
    tgtccaagtg ttctcattgt tcagttccca cctgtgagt agaacatgtg gtgtttggtt 780
    ttctttctct gcaatagttt gctcagagtg atgggtttcca gcttcgtcca tgtccctaca 840
    aaggacatga actcatcctt ttttatgact gcatagtatt ccgtgggtga tatgtgccac 900
30 attttcttaa tccagcttat catcgatgga catttgggtt ggttgcaagt ctttctact 960
    gtgaatagtg ccgcaataaa catacgtgtg catgtgtctt tatagcagca tgatttataa 1020
    tcctttgggt atataccca taatgggatg gctgggtcaa atggatttc tagttctaga 1080
    tccttgagga atcaccacac tgcctccac aatgggtgaa ctagtttaca cttccagcaa 1140
    cagtgtaaaa gtgttctggt gctggagagg atgtggacag cagttatttt ttatgaaaa 1200
35 tagtatcact gaacaagcag acagtttagt aaggatgctg cagggaagcct gcaggccaca 1260
    cagccatttc tctcgaagac tccgggtttt tctgtgcat cttttgaaac tctagctcca 1320
    attatagcat gtacagtgga tcaaggttct tcttcattaa ggttcaagtt ctatattgaa 1380
    ataagtttat gtaacagaaa caaaaatttc ttgtacacac aacttgctct gggatttggg 1440
    ggaaagtgtc ctgagctgg cggcacactg gtcagccctc tgggacagga tacctctggc 1500
40 ccattggcat ggggcgctgg gcttgggctt gagggtcaca cagtgcacca tgcctagctt 1560
    cctgtggata ggtctgggt ctcggatcat gctgaggacc acagctgcca tgcctggtaa 1620
    gggcaccacg tggctcagag ggggcgaggt tcccagcccc agctttctta ccgtcttcag 1680
    ttatttttcc ctaagagtct gagaagtggg gccgcgctg atggccttcg ttcgtcttca 1740
    gctggcacag aattgcacaa gctgatggta aacactgagt acttataatg aatgaggaat 1800
45 tgctgtagca gtttaactgt gagagctcgt ctgttggaaa gaaatttaag tttttcattt 1860
    aaccgctttg gagaatgtta ctttatttat ggctgtgtaa attgtttgac attcagctcc 1920
    tcgtagacag atactacgta aaaagtgtaa agttaacctt gctgtgtatt ttcccttatt 1980
    ttag 1984

50 <210> 14
    <211> 1871
    <212> DNA
    <213> Homo sapiens

    <400> 14
55 gtgaggcccg tgccgtgtgt ctgtggggac ctccacagcc tgtgggcttt gcagttgagc 60
    ccccggtgtc ctgccccctg caccgcagcg ttgtctctgc caagtctctt ctctctgccg 120
    gtgctggatc cgcaagagca gaggcgcttg gccgtgcacc caggcctggg ggcgcagggg 180
    caccttcggg agggagtggt taccgtgcag gccctggttc tgcaagagac caccaggtt 240
60 acacacgtgg tgagtgcagg cggtagacct gctcctgctg ctctttggaa agtcaagagt 300
    ggcggtctct ggggccccag tgagaccccc aggagctgtg cacagggcct gcagggccga 360
    ggcgagcagc tcctccccag ggtgcacctg agcctgcgga gagcaggagc tgcctgagtga 420
    gctggcccac agcgttcgct cgggtcacgt tcctgctggt ggttgtttgg gatcggtggg 480
    agaatttggg tttgtgagt gctgtgtctt tgaaccacgg agatggctag gagggtgtt 540
65 cagagttgat ttttctgaat caaactaaaa tcaggcacag gggacctggc ctgagcacag 600

```

14 / 18

5 gggattgtcc aatgtggtcc ccctcaaggg cgtccacacag agccggtggg cttgttttaa 660
 agtgcgattt gacgagggac gagaaacctt gaaagctgta aagggaaacc tcagaaaatg 720
 tggccgccag ggggtggttc aggtgctttg ctgggctgtg tttgtgaaaa cccattttgga 780
 cccgccctcc aagtcacccc tccagggtcca ccctccaggg ccgccctggg ctgggggtat 840
 10 gcctggcggt ccttgtgccc cagcccggag cacagcaggc tgtgcacatt taaatccact 900
 aagattcact cggggggagc ccagggtcca agcaactgag ggctcaggag tccctagggt 960
 gctgagggga cagagcagac ggggaacgct gcttctgtgt ggcaagttcc tgaggggtgt 1020
 gggccaggag gtggctcaga gtgtatgtt ggtcccacc gggggcagaa ctctgtctct 1080
 gatgagtcgg cagccatgta acaggaaggg gtggccacag ggagctggga atgcaccagg 1140
 15 ggagctcgcc agctggccga ggtcccaggg ccaggccaca ggaagggcag ggggacgccc 1200
 ggggcccacag cagaggccgc aggaagggaa ggggatgcc aggccagagc agaggctacc 1260
 gggcacaggg gggctccctg agctgggtga gcgaggctca tgactcggcg agggaaacct 1320
 cttgacgtga agctgacgac tgggtgtgccc cagctcacag cccagccagg tccccgcgct 1380
 gagcaggaaac tcagaacctt cccctttgtc taaagcacag cagatgcctt cagggcatct 1440
 20 aggagaaaaac aggcataagc gttgagaaac gtcttaaaag aagggtggat ggtggcaatt 1500
 tcttgtccag atttttagtct gccccggacc acagatgagt ctataacggg atttgtgtgt 1560
 tgccatgggg acacatgaga tggaccatca cagaggccac tggggctgca cctcccatct 1620
 gagtctggc tgtcccgggt ccaggccagg ttcttgcatg ctcacctacc tgtcctgccc 1680
 gggagacagg gaaagcacc cgaagtctgg agcagggtg ggtccagggt cctcagagct 1740
 25 cctgccaggc ccagcaccct gctccaaatc accactctc tgggggtttc caaagcattt 1800
 aacaagggtg tcaggttacc tctgggtga cggccccgca tctgggggt gacattgccc 1860
 ctctgctta g 1871

<210> 15
 <211> 3801
 <212> DNA
 <213> Homo sapiens

30 <400> 15
 gtgagcgac ctggccggaa gtggagcctg tgccccgctg gggcagggtg tctgagagg 60
 ccgttgctc caccctctgt tccgtgtggg gcaggcgact gccaatccca aagggtcaga 120
 gggcacaggg tgcccctcgt cccatctggg gctgagcaga aatgcattt tctgtgggag 180
 tgagggtgct cacaacggga gcagttttct gtgctatttt ggtaaaaagg aatggtgac 240
 35 cagacctggg tgactctagg tctcttcaga aagcagctct gatccgaacc caagacgccc 300
 gggccctgct gggcgtagt ctctcaaac cgaacacagg gggcctgctg ggcctagatc 360
 cctctgaacc cgagaccctg gggccctgct gggcgtagt ctctccgaac ccagagact 420
 caggggccct tggggcgtga gtctctccgc tgtgagcccc acactccaag gctcatccac 480
 agtctacagg atgcatgag ttcatgatca cgtgtgaccc atcaggggac agggccatgg 540
 40 tgtggggggg gtctctaca aattctgggg tctgtttcc ccagagccc agagctcaag 600
 gccccgtctc aggtctagac acaaatgaat tgaagatgga cacagatgca gaaatctgtg 660
 ctgtttctt tatgaataaa aagtatcaac attccaggca gggcaagggt gctcacacct 720
 ataatcccag cactttggga gggcgagggt ggtggatcac ttgaggccag gagtgtgagg 780
 ccaacctaac caacatagtg aaattccatt tctactaaa aaatacaaaa attagctagg 840
 45 cctgggtggc cacgcctgta gtccccgcta tggggaggc tgaggcagga gaatcattt 900
 aaccaggag gcagagggtg cagtgcagg agatcacacc actgcactcc agcctgggca 960
 acagagttag acttcatctt aaaaaaaaaa aaaaaagtat cagcattcca aaaccatagt 1020
 ggacagggtg ttttttattc tgtccttcga taatatctac tgggtgctgtg ctgagggccg 1080
 gaactggggg tgccttcctc tgaaggacac accttcatgg gaagagaaat aagtgtgaa 1140
 50 tgggtgttaa accagagggt taaactgggg tccgtgctgt ctgagttaac agtccagatc 1200
 tggactttgc ctctttccag aatgctccct ggggtttgct tcatggggga gcagcagggt 1260
 tggacaccct cgtgagggg gagcagcagg tgcagacgcc ctcatgatgg gggagtggca 1320
 ggtgcagaca cctctgtgca tgggtcccag catgtccctg ttgcagctcc ctccccaca 1380
 ggatgccgggt ctctgtgct cccacagtc cctgcttccc tctcacagcc ttacctgggt 1440
 55 ctggcctcca ctggctttgt ctgcatgatt tccacattc ctgggctccc agcactctct 1500
 cgcctctccc aggcacctct gcagtgtgg ccataccagt cagctgtgaa ctgtccactg 1560
 cttattttgc tccccatgaa atgtatttt taggacaggc acccctgggt ccagcctctg 1620
 gcacagatc agtgaatgtt attgaaggac aaaggacaga caaacaatc aggaaaatg 1680
 gttctctcta aacacattg aaagccacag aggttagtgc aggatgggtg ggcatcagg 1740
 60 catcagatgt ggggtccaat ccagaatatt ctgtgctccc aaaggccact tggtcagagt 1800
 gtgtgcttgc agaggtggct ctaaaagctc agcagtgagg gcagtgttcc gccatctca 1860
 ggggtgaactc acatcctctg tgtctgaagt atacagcaga ggcttgaagg gcatctggga 1920
 gaagaaaaca ggcaaatga ttaagaaaag tgaagaaagg aagtggttaa gatgggaatt 1980
 65 tctttgtcca gatttttagt tcccaaacca cagctcagat ggtagaatgt ggtcagaact 2040
 gatggacaga acaatagaac aaaacggaag ccctatctct cagaaacgtg tgttaatgtg 2100
 gtatgtggca cagctgatgg aaaagagagt gtgtgtgtaa tttttttttc tgagaaaact 2160

15 / 18

gactggaagc aaataagttg tgtctttaca gcatatacca gagcagattc taggtagaag 2220
aggagacaca tgcaaacac accagcaaca gaaataaaac aaaagactca aaggggaagg 2280
agggtgaacgt tccctgggtt ggtgttgggg aaggacacac agggaggcgg atgaaaccag 2340
tgaggcaacg ggcattgctt tcaatgcaga gaaactcagc ttgcctgagc cacagtgaag 2400
5 atggccattc cctggagcgt ttgtgcacgt gatttattta aggcgccttg tgaggctctg 2460
cacattcatc ctctcacttt gtctctctaa ccacctgaga ggttagaggag gaaaggctcc 2520
aggggagcag cgcgccttgg tcacccagct ggcaaaagggc atgcattgatt gcagcctggc 2580
ctctctgctc ggggcccctg ctctgcccga ggaccccaaca caagtcagac ccataggctc 2640
10 aggggtgagcc ggagcccaag gtcgtgttgg ggtgggtgtg gaaagaagaa atggagctct 2700
gatgcacact tgggaagggtc ctaccagcag cgtcaaaagaa atgcattgta aactgacagc 2760
gagaccatc cctcaaaagaa acgcacgtga aactgatggc gagacctgtc cccatccctc 2820
atgctggctc cttttctggg cttgcccaaga gccagcatca ggttgaggca agctggaaag 2880
acttttctgg aaagcagctt gtttgcattg aagtcctcac aatgtcctgt gtcttccag 2940
taattccact tctgaagtga ccagacatta tcacgggtct tatttaccat ttccagtgtt 3000
15 ccaggcaggg ggacttgcca cagcaagtca cgaacctgac caatacagg gctaaggaga 3060
tattatgcat cacaaaactt gctctgccat taaacatttt tcaagaagaa ttggaagaat 3120
gtttaatggc acaaacgtt tatttcaatg tagcagtggt caaagctgga tgtaaaagaa 3180
cacacccag gagcctgccc tgaatgtcat gtgtgttcat ctttgacat ggacatcat 3240
gggcagtgag tgggtgtgag gccctggagg acatcggtgg gatgcctcca tctgcccct 3300
20 ctggagacac catgtgtgcc acgtgcactc actggagccc tgtttagctg gtgccacctg 3360
gctcttccat cctcgagatt caaacacagt gagattcccc acgcccactc cagtgtctc 3420
ccacaaaaaa cctgagtcac acctgtgttc actcgaggga cgcccgggag ccagggtccc 3480
acagcttatt atgtgttttt ggtcgagtta tgtgcagatc tcatcagggc agatgatgag 3540
25 tgcacaaaaa cggccgtgag aggtttggat acactcaaca tcaatagcca ggtcctggtg 3600
gagtttggtc atgcagagtc tggatggcat gtagcatttg gagtccattg agtgagcacc 3660
cagccccctc gggctgcagc gcatgcccga ggaggagcaa ggaagcggga ggaagcgagg 3720
aggctctttg gagcaagctt tgcaggaggg ggctgggtgt ggggcaggca cctgtgtctg 3780
acattcccc cttgtgtctca g 3801

30 <210> 16
<211> 880
<212> DNA
<213> Homo sapiens

35 <400> 16
gtgagcaggc tgatgggtcag cacagagttc agagtccagg aggtgtgtgc gcaagtatgt 60
gtgtgtgtgt gtgcgcgcgt gcctgcaagg ctgatgggtg ctggctgcac gtaagagtgc 120
acatgtacgc atatacagct gagcacatca atgtgtgcat gtgtgtacat gaaggcatgg 180
40 cagtgtgtgc acaggtgtgc aagggcacaa gtgtgtgcac atgcgaatgc acacctgaca 240
tgcatgtgtg ttcgtgcaca gtcgtgtggg cattcacgtg aggtgtcatg gtgtgggtgt 300
gcagtgtgag tagcatgtgt gcacataaca tgtattgagg ggtcctcgtg ttccccccg 360
taggtcctca gcaccagtg cactccttac aggatgagac ggggtcccg gccttggtg 420
gctgaggctc tgaagctgca gccctgaggg cattgtccca tctgggcac cgcgtccact 480
50 cctctcctg tgggcttctg tgtccactcc cctctcctg tgggcattta catccactcc 540
actcctctc tctgtgggc atccgcgtcc actccccctc tctgtgggca tctgcgtcca 600
cctccccct ctgtgggcat ttgcgtccac tccctctcct ggttccctcc tgtcttggtc 660
gagcctcggg ggcaggcaga tgacacagag tcttgactcg cccagggtgg ttcgcagctg 720
ccgggtgagg gccaggccgg atttcaactg gaagagggat agtttcttgt caaaatgttc 780
ctctttctg ttccatctga atggatgata aagcaaaaag taaaacttta aaatcccaga 840
50 gaggtttcta cgttttctca ctcttctctg gcgactctag 880

<210> 17
<211> 3186
<212> DNA
55 <213> Homo sapiens

<400> 17
gtgagccggc accaaggggt gcaggcccag cctccaggga ccctccgcgc tctgctcacc 60
tctgacccgg ggtttcacct tggaaactcct ggggttttag ggcaaggaaat gtcttacgtt 120
60 ttcatgtgtg ctgctgcctg tgcacagttc tgttcgcgtg gctctgtgca aagcacctgt 180
tctccatctc tgggtagtgg taggagccgg tgtggcccca ggtgtcccca ctgtgctgt 240
gcactggccg tgggacgtca tggaggccat cccaggggcag caggggcatg gggtaaaag 300
atgtttatgg ggagctctag cagaggaggc tgggaagggt tctgaacagt agatgggaga 360
tcagatgcc ggaggatttg ggggtctcagc aagaggggcc gaggtgggtg cagggtgagg 420
65 tgcgtggccc caccgccggg aagggtgcagc agagctgtgg ctccccacac agcccgcca 480

16 / 18

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

```

gcacctgtgc tctgggcatg gctgtgtccc tggaaacgttc cctgtcctgg ctgggtcagg 540
gggtgccctg ccaagaatcg acaactttat cacagaggga agggccaatc tgtggaggcc 600
acagggccag cttctgcctg gagttagggc aggtggtggc acaagcctcg gggctgtacc 660
aaagggcagt cgggcaccac agggccgggc ctccacctca acaggcctcc cgagccactg 720
ggagctgaat gccagggagg cgaagccctc gcccatgag ggctgagaag gagtgtgagc 780
atttgtgtta cccagggccg aggtgtgcgc aattaccgtg cacacttgat gtgaaatgag 840
gtcgtcgtct atcgtggaaa cccagcaagg gctcacggga gagtcttcca ttacaaggtc 900
gtaccatgaa aatgggtttt aaccgagtg cttgcgcctt catgctctgg cagggagggc 960
agagccacag ctgcatgtta ccgcttttgc accagctcca gaggcttggg accaggctgt 1020
ctcagttcca ggggtgcgtc ggctcagacc gccctcctct ctgctctctc tctctgcctc 1080
aaatcttccc tcgtttgcat ctccctgacg cgtgcctggg ccctcgtgca agctgcttga 1140
ctcctttccg gaaacccttg ggggtgtgtg gatacagggt ccactgagga ctggagggtg 1200
ctgacactgt ggttgacccc aggggtccagc tggcgtgctt ggggctcctc tgggcatga 1260
tgaggtcaga ggagttttcc caggtgaaaa ctctgggaa actccccagg ccatgtgacc 1320
tgccacctgc tctcccataa ttcagctcag tctgtctctc atttcccccac cagggtctct 1380
agctccgagg agctccccta gagggcctgg gctcaggcca gggcggtgta gtttccccac 1440
ccatgtgggg acccttgggt agtcgttga ttgggtagcc ctgaggaggc cgagatgcga 1500
tggggccagg gccgtttcca aacacagagt caggcacgtg gaaggccagc gaatccccct 1560
cccccgaggc aggagtggga gaacggagag ctgggccccg atttcacggc agccaggctg 1620
cagtgggcga ggctgtgtgt gtccacgtgg cgctgggggc ggggtctgat tcaaaccgc 1680
tggggctcgg ccttctctgg ccgtgtgtgg cgcgctcca cagggtctg ggggtggacg 1740
cccgaactct agcaggtggc tatttctccc ttggaaagag agccccccac ccatgctagg 1800
tgtttccctc ctgggtcagg agcgtggccg tgtggcaacc ccgggacctt aggtcttatt 1860
atttgtttaa aaacattctg ggcctggctt ccgttgttgc taaatgggga aaagacatcc 1920
cacctcagca gaggttactg gaggtgaaa ccgggggtgt ggcttgactg gtgtgatctc 1980
aggtcattcc agaagtggct caggaagtca gtgagaccag gtacatgggg ggctcaggca 2040
gtgggtgaga tgaggtacac ggggggctca ggcagtgggt gagggcaggt acatgggggg 2100
ctcaggcact ggggtgagat aggtacacgg ggggctcagg cagaggggtc gaccagggtac 2160
acgggggctc tgatcacacg cacatatgag cacatgtgca catgtgctgt ttcatggtag 2220
ccaggtctgt gcacacctgc cccaaagtcc caggaagctg agaggccaaa gatggaggct 2280
gacagggctg gcgcggtggc tcacacctgt agtcccagca ctttgggagg ccgaggcgag 2340
aggatccctt gagcccagga gttaagacc agcctgagca acatagtaga accccatctc 2400
tatgaaaaat aaaaaaataa attagctgaa catgggtggg tgcgcctgta gttccaatac 2460
ttgggaggct gaagtgggag gatcacttga gcccaggagg tggaaagtgc agtgagtga 2520
gattgcacca ctgtactgca gcctgggtga cagagtga gcccattcca acaacaacaa 2580
agaagactga caaatgcagt ttcttggaaa gaaacattta gtaggaaact aacctacaca 2640
cagaagccaa gtcgggtgtc cgggtgtcag gagatgagat gatgggtcct cacaccatca 2700
ccccagacct aggggtttatg caccacaggg gcgggtggct cagaagggat gcgcaggagc 2760
ttgatatacy atgacatcaa ggttgtctga cgaagggcag gattcatgat aagtaacctg 2820
tggtaacaaa ggaacaatgg ataaactgga aaccttagag gccttcccg aacaggggct 2880
aatcagaagc cagcatgggg ggctggcatc caggatggag ctgcttcagc ctccacatgc 2940
gtgttcatac agatgggtgca cagaaacgca gtgtacctgt gcacacacag acacgcagct 3000
actgcacac acaagcacac acacagacat gcatgcatgc atccgtgtgt gtgcacctgt 3060
gcccagagg aaacccatgc atgtgcattc atgcacgcac acaggcaccg gtggggccat 3120
gcccacaccc acgagcaccg tctgattagg aggcctttcc tctgacgtg tccgccatcc 3180
tctcag
  
```

<210> 18
 <211> 781
 <212> DNA
 <213> Homo sapiens

<400> 18
 55
 60

```

gtatgtgcag gtgcctggcc tcagtggcag cagtgcctgc ctgctgggtg tagtgtgtca 60
ggagactgag tgaatctggg cttaggaagt tcttaccctt tttcgcatac ggaagtgggt 120
taaccacacc actgtcaggc tcgtctgccc gccctctcgt ggggtgagca gagcacctga 180
tggaaaggac aggagctgtc tgggagctgc catccttccc accttgctct gcctggggaa 240
gcgctggggg gcctgtgtct tectgtttgc cccatgggtg gatttggggg gcctggcctc 300
tctctgtttg cctgtgtgtg gattggggtg tctcccgctc atggcactta gggcccttgt 360
gcaaacccag gccaaagggt taggaggagg ccaggccacg gctaccccac cctctcagg 420
agcagaggcc gcgtatcacc acgacagagc cccgcgccgt cctctgcttc ccagtcaccg 480
tctctgccc ctggacactt tgtccagcat caggagggtt tctgatccgt ctgaaattca 540
agccatgtcg aacctgcggg cctgagctta acagcttcta ctttctgttc tttctgtgtt 600
  
```

17 / 18

gtggaaattt cacctggaga agccgaagaa aacatttctg tctgtactcc tgcggtgctt 660
gggtcgggac agccagagat ggagccaccc cgcagaccgt cgggtgtggg cagctttccg 720
gtgtctctctg ggaggggagc tgggctgggc ctgtgactcc tcagcctctg ttttccccc 780
g 781

5
<210> 19
<211> 536
<212> DNA
<213> Homo sapiens

10
<400> 19
gcaagtgtgg gtggaggcca gtgcgggccc cacctgccca ggggtcatcc ttgaacgccc 60
tgtgtggggc gagcagccctc agatgctgct gaagtgcaga cggcccccggg cctgaccctg 120
ggggcctgga gccacgctgg cagccctatg tgattaaacg ctggtgtccc caggccacgg 180
15 agcctggcag ggtccccaac ttcttgaacc cctgtctccc atctcagggg cgatggctcc 240
ccaagcttgg gagccttctg acccctgacc tgtgtcctct cacagcctct tccctgggtg 300
ctgccctgag ctccctggggc cctgagcaag ttctctcccc gcccgcgcgc tccagcgtca 360
ctgggctgcc tgtctgctcg ccccggtgga ggggtgtctg tcccttcaact gaggttccca 420
ccagccaggg ccacgaggtg caggccctgc ctgcccggcc acccacacgt cctaggaggg 480
20 ttggaggatg ccacctctgg cctcttctg aacggagtct gattttggcc ccgcag 536

<210> 20
<211> 3179
<212> DNA
25 <213> Homo sapiens

<400> 20
atctcatgtt tgaatcctaa tgtgcactgc atagacacca ctgtatgcaa ttacagaagc 60
ctgtgagtga acggggtggt ggtcagtgcg gggccatggc ctggctgtgc atttacggaa 120
30 gtctatgagt gaatggggtt gtggtcagtg cgggcccagc gcctggctgg gcctggaggg 180
ttctctatgc tgtgaggcag gaggggaagg agggtagggg atagacagtg ggagccccc 240
ccctggaaga cataacagta agtccaggcc cgaaggggcag cagggatgct gggggcccag 300
cttgggcggc ggggatgatg gagggcctgg ccagggtggc agggatgatg gggggcccag 360
ctggggtggc aggggtgatg gggggggctg gtctgggtgg cggggaagat ggggaagcct 420
35 ggctgggccc cctcctcccc tgccctccac ctgcagcctg ggatccggat gtgttctcct 480
gggtcacatc ctctggggcca tcagctttca tggagggtgg gggcaggggc atgacacat 540
cctgtataaa atccaggagt cctcctctg aacgccccaa ctcaggttga aagtcacatt 600
ccgctctgg ccattctctt aagagtagac caggattctg atctctgaag ggtgggtagg 660
gtggggcagt ggagggtgtg gacacaggag gcttcagggt ggggctgtgt atgtctcttc 720
40 atctctctat catctcccag tctcatctct catctctta tcatctccca gtctcatctg 780
tcttctctct atctcccagt ctcatctgtc atcctcttac catctcccag tctcatctct 840
tatectctta tctcctagtc tcatccagac ttacctccca gggcgggtgc caggctcgca 900
gtggagctgg acatacgtcc ttctcaggc agaaggaaact ggaaggattg cagagaacag 960
gagggggcgc tcagagggac gcagctctgg ggtgaagaaa cagccctcc tcagaagtgt 1020
45 gcttgggcca cagaaaaccg agggccctgc gtgagtggct ccagagcctt ccagcaggtc 1080
cctggtgggg ccttatggta tggccgggtc ctactgagtg cactctggac agggcttctg 1140
gtttgagtgc agcccgagc tgctgtgtgt cggggtgggg gcttatggcc actggatatg 1200
gcgtcattta ttgctgctgc ttcagagaat gtctgagtga ccgagcctaa tgtgtatggt 1260
gggcccgaag ccacagactg tgcgtaaat gcactctggt gcctggagcc cccgtatagg 1320
50 agctgtgagg aaggaggggc tcttggcagc cggcctgggg gcgcctttgc cctgcaaaact 1380
ggaaggaggc gggcccgggc gccgtgggag gacgacctca agtgagagggt tggacagaa 1440
agggcgggga ctccccagga gcagaggccg ctgctcaggc acacctgggt ttgaatcaca 1500
gaccaacagg tcaggccatt gttcagctat ccattctcta caaagctcca gattcctgtt 1560
tctccgggtg ttttttgttg aaattttact caggattact tatatttttt gctaaagtat 1620
55 tagaccctta aaaaaggat ttgctttgat atggcttaac tactaagca cctactttat 1680
ttgtctgttt ttatttatta ttattattat tattagagat ggtgtctact ctgtcaccca 1740
ggttgttagt gcagtggcac agtcaggtc cgtgttagcc gcaaaacccc aggtcgaagt 1800
gatcctccgg cctcagcttc ccagagtgtt gggattacag gtgtgagcca ctgcccttgc 1860
ctggcacttt taaaaaccac tatgtaaggc cagggtccagt ggcttccaca cctgtcatcc 1920
60 cagtagtttg ggaagccgag gcagaaggat tgtctgaggc caggagtttg agaccagcat 1980
gggtaacata gggagacccc atctctacaa aaatgcaaa aagttatccg ggcgtggggg 2040
ccagcatctg tagtcccagc tgctcgggag gctgagtggt aggatcgctt gagcccgga 2100
ggtcattggt gcagtgtgct gtgattgtac catcgactc cagcctgggc aacagagtga 2160
gaccctgtct caaaaaaaaa aaaaaaaaaa gaaggagaag gagaagagaa gaagaaggaa 2220
65 gaaggaaaga gaagaagaag gaagaaggaa gaaagaagga gaaggaggcc tgcagggtgc 2280

18 / 18

taggtagact gtcaaatctc agagcaaat gaaaataaca aagtttttaa gggaaagaaa 2340
aaccaccagct ctttggactt ccttaggcct gaacttcac tcaagcagct tccttccaca 2400
gacaagcgtg tatggagcga gtgagttcaa agcagaaagg gaggagaagc aggcaagggt 2460
ggaggctgtg ggtgacacca gccaggaccc ctgaaaggga gtggttgttt tcctgcctca 2520
5 gccccacgct cctgccggtc ctgcacctgc tgtaacctgc gatgttggtg ccagggtgccc 2580
acctgggaag gatgctgtgc agggggcttg ccaaaactttg gtgggtttca gaagccccag 2640
gcacttgtgg caggcacaat tacagccct ccccaaagat gcccacgtcc ttctcctgga 2700
acctgtgaat gtgtcacccg caaggcagag gctggtgaag gctgcaggtg gaatcacggc 2760
tgccagtcag ccgatcttaa ggtcatcctg gattatctgg tgggcctgat atggccacaa 2820
10 gggtccttag aagtgagaga gggaggcagg ggagagtcag agaggggacg tgagaaggac 2880
cactggccac tgctggcttt gagatggagg aggggggtccc cagccaagga atgggggcag 2940
ccgctccatg ctggaaaagc aagcaatcct ccccggtcct gagggcacac ggccctgccc 3000
acgcctcgat ttcaggccag tgggacctgt ttcagctttc cggcctccag agctgtaaga 3060
tgatgcgttt gtgttcagcc actaagctgc agtgattcgt cacagcagca aatggaatag 3120
15 cagtacaggg aaatgaatac agggacagtt ctacagtgta ctctcagccc acccctggg 3179